

From: Fredman, Jeffrey
Sent: Tuesday, January 25, 2005 6:05 AM
To: STIC-Biotech/ChemLib
Cc: Dunston, Jennifer
Subject: FW: Sequence Search 10/826523

PLEASE RUSH.

I Approve.

Jeff Fredman

-----Original Message-----

From: Dunston, Jennifer
Sent: Monday, January 24, 2005 12:46 PM
To: Fredman, Jeffrey
Subject: Sequence Search 10/826523

Jeff,
Please RUSH this search. The length of SEQ ID NO: 40 is 707 nucleotides. The length of SEQ ID NO: 41 is 3662 nucleotides. The sequences are related in that SEQ ID NO: 41 comprises the sequence of SEQ ID NO: 40.
Jenn

Please do a sequence search for the nucleic acid sequence of SEQ ID NO: 40 and SEQ ID NO: 41 against the commercial and interference nucleotide databases.
Thank you.

Jennifer Dunston, Ph.D.
USPTO Art Unit 1636
REM 2B76
Mailbox: REM 2C70
(571) 272-2916

STAFF USE ONLY

Searcher: Jan
Searcher Phone: 2- 2504
Date Searcher Picked up: 1/25/05
Date Completed: 1/27/05
Searcher Prep/Rev. Time: 10
Online Time: 10

Type of Search

NA Sequence: # ✓
AA Sequence: #
Structure: #
Bibliographic:
Litigation:
Patent Family:
Other:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other(Specify):

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:27:35 ; Search time 15598.5 Seconds
(without alignments)
11102.026 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662
Sequence: 1 ctaaatgtgaagcgttaata.....attccccgaagaagtcacc 3662

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2654.8	72.8	3637	6	AX592968 Sequence
C 2	2656.4	72.5	3637	6	AX592967 Sequence
C 3	2650.6	72.4	3626	12	ASPGREB1 Y09373 Artificial
4	2644.8	72.2	3624	12	XXU43957
5	2639	72.1	3681	6	A20702
6	2639	72.1	3681	6	115651
7	2623.6	71.6	5314	6	A20700
8	2623.6	71.6	5314	6	115649
C 9	2610.8	71.3	3552	12	AF171227
C 10	2609.4	71.3	4001	6	AX306330
11	2558.8	69.9	5540	12	AF179627
C 12	2550.6	69.9	3633	12	ASPGREB2 Y09374 Artificial
C 13	2537.8	68.8	3918	6	CQ768840
C 14	2517.8	68.8	3918	6	CQ802124
15	2514.6	67.9	4141	12	XXU5129
C 16	2485.8	67.9	4141	12	AY190524
17	2484.6	67.6	5277	6	AX306326
18	2474.2	67.4	5277	6	BD244458
19	2469.4	67.4	5277	6	AR340565

20	2451.8	67.0	3928	6	AX591027	AX591027 Sequence
21	2451.8	67.0	3928	6	AX591180	AX591180 Sequence
22	2451.8	67.0	3928	6	AX717603	AX717603 Sequence
23	2449.8	66.9	7933	12	AF359244	AF359244 Integrati
24	2448.8	66.9	3956	6	A84776	A84776 Sequence 21
25	2448.8	66.9	3956	6	AR300179	AR300179 Sequence
26	2448.8	66.9	3956	6	BD058066	BD058066 Method fo
27	2446.8	66.8	3715	6	AX306328	AX306328 Sequence
28	2434	66.5	3970	12	CVE403983	AJ403983 Cloning v
C 29	2420.6	66.1	3274	12	XXU43956	U43956 Expression
C 30	2414.6	65.9	3978	12	CVU41513	U41513 3' terminal
C 31	2401	65.4	3733	12	AF146395	AF146395 Cloning v
C 32	2396	65.1	3307	12	SYNPEBXC	D50399 Unidentifc
C 33	2382.6	65.1	4591	12	SYNPEZ18A	M74186 Cloning vec
C 34	2380.4	65.0	6624	12	CVU37458	U37458 Yeast CUP1
C 35	2372.2	64.8	4239	12	AVE401048	AJ401048 Artificial
36	2359	64.4	5781	12	CVU37457	U37457 Yeast CUP1
37	2357.4	64.4	4051	12	XXU5126	U35126 Yeast CUP1
C 38	2351	64.2	5275	12	CVU37456	U37456 Yeast CUP1
C 39	2349.8	64.2	3230	12	AY663843	AY663843 Cloning v
C 40	2340.2	63.9	3230	12	AY663844	AY663844 CFP Integ
C 41	2333.4	63.7	5436	12	AY303167	AY303167 YFP Integ
C 42	2332.8	63.7	5436	12	AY303167	AY303167 YFP Integ
C 43	2328.6	63.6	3204	12	AY663845	AY663845 Cloning v
C 44	2327.2	63.5	5385	12	IVE250107	AJ250107 Integrati
C 45	2322.6	63.4	5445	12	IVE306911	AJ306911 Integrati

ALIGNMENTS

RESULT 1
LOCUS AX592968/c 3637 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 4 from Patent WO02079468.
ACCESSION AX592968
VERSION AX592968.1 GI:27950812
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1 Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
AUTHORS
TITLE A method of increasing complementarity in a heteroduplex
JOURNAL Patent: WO 02079468-A 4 10-OCT-2002;
Large Scale Biology Corporation (US)
FEATURES
source
1..3637
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match	Score	72.8%	Score	2664.8	DB	6	Length	3637
Best Local Similarity	83.8%	Pred.	No. 0					
Matches 3070	Conservative	0	Mismatches	567	Indels	25	Gaps	4
1	CTAAATGTGAAGCGTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC	60						
3637	CTAAATGTGAAGCGTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC	3578						
61	ATTTTAAACCAATAGCGCCGAATATCCCTTAATATCAAAAGATAGACCGA	120						
3577	ATTTTAAACCAATAGCGCCGAATATCCCTTAATATCAAAAGATAGACCGA	3518						
121	GATGGGTGAGTGTTCAGTTTGAACCAAGTCCCTTAATTAAGAACGTC	180						
3517	GATGGGTGAGTGTTCAGTTTGAACCAAGTCCCTTAATTAAGAACGTC	3458						
181	CAACGTAAAGGCGCAAAACCGTCTATCAGGCGATGCGCCACTAGTGAACCATCACC	240						
3457	CAACGTAAAGGCGCAAAACCGTCTATCAGGCGATGCGCCACTAGTGAACCATCACC	3398						

QY 241 CTAATCAAGTTTTGGGGTCGAGGTCCGTTAAAGCACTAAATCGGAACCTTAAAGGAG 300
| | | | |
Db 3397 CTAATCAAGTTTTGGGGTCGAGGTCCGTTAAAGCACTAAATCGGAACCTTAAAGGAG 3338
| | | | |
QY 301 CCCCCGATTTAGAGCTTGAACGGGAAAAGCCGGCGAAGCGGGGGAAGAAAGGAAAGAA 360
| | | | |
Db 3337 CCCCCGATTTAGAGCTTGAACGGGAAAAGCCGGCGAAGCGGGGGAAGAAAGGAAAGAA 3278
| | | | |
QY 361 AACGAAAGAGAGCGGCGCTGAGGGCGCTGGCAAGGTAGCGGTACAGCTGCGGTAAACCA 420
| | | | |
Db 3277 AACGAAAGAGAGCGGCGCTGAGGGCGCTGGCAAGGTAGCGGTACAGCTGCGGTAAACCA 3218
| | | | |
QY 421 CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCATTGCGCATTCAGCTCG 480
| | | | |
Db 3217 CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCATTGCGCATTCAGCTCG 3158
| | | | |
QY 481 CACCTGTTGGGAAGGGCGATCGGTGCGGGCCCTTTCGCTATTACGCCAGCTGGCGAAAG 540
| | | | |
Db 3157 CACCTGTTGGGAAGGGCGATCGGTGCGGGCCCTTTCGCTATTACGCCAGCTGGCGAAAG 3098
| | | | |
QY 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAAACGCGAGGGTTTTCCAGTCAGACGTTG 600
| | | | |
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAAACGCGAGGGTTTTCCAGTCAGACGTTG 3038
| | | | |
QY 601 TAAACGACGCGCAGTGAACGCGCTCGTTCATTCAAGTTTTTGAACCCGCGAGGACGG 660
| | | | |
Db 3037 TAAACGACGCGCAGTGAACGCGCTCGTTCATTCAAGTTTTTGAACCCGCGAGGACGG 2978
| | | | |
QY 661 GCACAGCTCGCGGTGCAAAATGTGTTTTTACGCGTATGAGACAGATGATGCTCGACAC 720
| | | | |
Db 2977 CGCGGCTGCGG-----CGGCTGAGAACTAGTGATGATCCCC--CGG 2939
| | | | |
QY 721 GCTCGAAGAACGCGAGCTAGATTAAACCTAGAAAGATTAATCATTTGTGACGTAACTTAA 780
| | | | |
Db 2938 GCTCGAAGAAATCTTATTTGTATGTTCATTCATCCAGTGTATCCGACGACAGTTA 2879
| | | | |
QY 781 AGATAATCATGCGTAAATTTGACGCGATGGATCTGTAAATACGACTCATTTAGGCGCAAT 840
| | | | |
Db 2878 CAATCTCAAGAGGAGCATGTGTGTCAGCTTTTGTGGGATCTTTGAAAGGCGAGATT 2819
| | | | |
QY 841 TGGGTACCGGGCCCCCCTCGAAGTTCGACGGTATCCATTAAGCTTGAATTCCTCGC 900
| | | | |
Db 2818 GTGTGACAGAGTAATGTTGTCTGTGTAAGGACAGGGCCATCGCAATGGAGTATTTT 2759
| | | | |
QY 901 AGCCCGGGGATCACTAGTTCTAGAGCGGCGCCACCGCGGTGAGCTCCAGCTTTTGT 960
| | | | |
Db 2758 GTTGATATAGTGTCTGCTAGTTGAACGGAATCATCTTCAATGTTGTGCGCAATTTGAAG 2699
| | | | |
QY 961 TCCCTTAGAGAGGTTAATTAATGCCATGCGTCAATTTTACGACGACTATCTTTCTAG 1020
| | | | |
Db 2698 TAGCTTAGATTCCATCTTTTGTGTTGTGTCGCGGTGATGATACATTTGTGAG--TTATAG 2640
| | | | |
QY 1021 GGTAAATCTAGTGCATCAGGATCATATCGTCCGGTCTTTTTCCGGCTCAGTACGCGC 1080
| | | | |
Db 2639 TTGTACTCGAGTTTGTGTCCGAGAAATGTTTCCATCTTCTTAAATCAATACCTTTTAAAC 2580
| | | | |
QY 1081 CAAGCTGCGCTTATCTGGGCATCTGGGAGAGAGAACCCGTGCTTTTCCCGGAGGTTG 1140
| | | | |
Db 2579 TCGATACGATTAAACAAGGGTATCACCTTCAACTTACGACAGCGGTCTTTGATGTTCC 2520
| | | | |
QY 1141 AAGCGGATAGGAAGATTGCCGAGATGACGCTGCTGCAATGACGTTGACGGAAGAAC 1200
| | | | |
Db 2519 CCGTCACTTTTGAAGATATAGTGCCTTCTGTACATTAACCTTCCGGGACATGCACTCTTG 2460
| | | | |
QY 1201 GCACTGTTAACATGATTCGGGAAGTGTGGCCATGCAAGCTTTTAAACGAGTAACTGT 1260
| | | | |
Db 2459 AAAAAGTACATGCGGTTTCAATATGATCCGATTAACGGGAAAAGCATTTGAACCATTAAGAG 2400
| | | | |
QY 1261 TCGTTAGCGCACCTGAGGATACCAAGTTGTCGCGGCTTTTCCGACACAGTTCCGAGATG 1320
| | | | |
Db 2399 AAAAGTAGTACAAGGTTGGCCATGGAACAGGATGTTTTCAGTAGTGCAAAATTAATTTA 2340
| | | | |
QY 1321 TCAGCCGAAGGCAATCAGCAACCGGAACAATACCGGCGACAGCCGGAATCTCGTCCG 1380
| | | | |

Db 2339 AGGTAAGCTTTCCGATGTAGCATCACTTCAACCTTCCATGACAGAAAATTTGTGC 2280
| | | | |
QY 1381 GTGTGCAGATTAAATGACAGGGGTGCGCGCTGGGAATTAATCTACAGAGACGGGTATC 1440
| | | | |
Db 2279 CCAATTAACATCACATTAATTAATCAACAAGATTGGGACAACTCAGATGAAGAAATCTTCTC 2220
| | | | |
QY 1441 CTGGCTGATAGCCGCAAGAAATGGAATATGATACCCGTGAGTTACCCGGCGGCGGCTT 1500
| | | | |
Db 2219 C---CTTTATCTCATCGGTACCAAGCTTTTGTCTCTTTAGTAGAGGTTAATTTGCGGCTT 2163
| | | | |
QY 1501 GCGGTAAATCAATGTATAGCTGTTTCTGTGTGAATTTGTATCCGCTCAATTCACAA 1560
| | | | |
Db 2162 GCGGTATCATATGTATAGCTGTTTCTGTGTGAATTTGTATCCGCTCAATTCACAA 2103
| | | | |
QY 1561 CAACATTAACGAGCCGGAAGCATTAAGTGTAAAGCTTGGGGTCCCTAATGAGTAACT 1620
| | | | |
Db 2102 CAACATTAACGAGCCGGAAGCATTAAGTGTAAAGCTTGGGGTCCCTAATGAGTAACT 2043
| | | | |
QY 1621 CACATTAATTTGCGTTCGCTCACTGCGCGCTTTCAGATGCGGAAACCTGTGCGACGT 1680
| | | | |
Db 2042 CACATTAATTTGCGTTCGCTCACTGCGCGCTTTCAGATGCGGAAACCTGTGCGACGT 1983
| | | | |
QY 1681 GCATTATATGATGCGCAACGCGCGGAGAGAGCGGTTTGCATTTGGCGCTCTTCCGC 1740
| | | | |
Db 1982 GCATTATATGATGCGCAACGCGCGGAGAGAGCGGTTTGCATTTGGCGCTCTTCCGC 1923
| | | | |
QY 1741 TTCTCGCTCACTGACTGCTGCGTGGGTGTTGGCGTGGCGGAGCGGTATCAGCTCA 1800
| | | | |
Db 1922 TTCTCGCTCACTGACTGCTGCGTGGGTGTTGGCGTGGCGGAGCGGTATCAGCTCA 1863
| | | | |
QY 1801 CTCGAAGCGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAAATGTG 1860
| | | | |
Db 1862 CTCGAAGCGGTAAATACGGTTATCCACAGAAATCAGGGATTAACGAGGAAAGAAATGTG 1803
| | | | |
QY 1861 AGCAAAAGCCGCAAAAGCCAGAGAACCGTAAAGCCGCGTTGCTGCGGCTTTTCA 1920
| | | | |
Db 1802 AGCAAAAGCCGCAAAAGCCAGAGAACCGTAAAGCCGCGTTGCTGCGGCTTTTCA 1743
| | | | |
QY 1921 TAGGCTCGCCCCCTGACAGCATCAAAAATGACGCTCAAGTCAAGAGTGGCGAA 1980
| | | | |
Db 1742 TAGGCTCGCCCCCTGACAGCATCAAAAATGACGCTCAAGTCAAGAGTGGCGAA 1683
| | | | |
QY 1981 CCGGACAGGACTATTAAGATACCAAGCGTTTTCCCTTGAAGCTCCCTGTCGCTCTCC 2040
| | | | |
Db 1682 CCGGACAGGACTATTAAGATACCAAGCGTTTTCCCTTGAAGCTCCCTGTCGCTCTCC 1623
| | | | |
QY 2041 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAACGTTGCG 2100
| | | | |
Db 1622 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAACGTTGCG 1563
| | | | |
QY 2101 GCTTTCTATAGCTACGCTGTAGGTATCTCAGTTGCGGTAGTGTGCTTCCAGACT 2160
| | | | |
Db 1562 GCTTTCTATAGCTACGCTGTAGGTATCTCAGTTGCGGTAGTGTGCTTCCAGACT 1503
| | | | |
QY 2161 GGGGCTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGAATCTATCG 2220
| | | | |
Db 1502 GGGGCTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGAATCTATCG 1443
| | | | |
QY 2221 TCTTGAATCAACCCGCTTAAAGACAGACTTATCGCACTGCGAGACGCACTGGTAAACAG 2280
| | | | |
Db 1442 TCTTGAATCAACCCGCTTAAAGACAGACTTATCGCACTGCGAGACGCACTGGTAAACAG 1383
| | | | |
QY 2281 GATTAGCAAGCGAGGTATGTAGCGGTGCTAACAAGTCTTGAAGTGTAGGCTTAATACTA 2340
| | | | |
Db 1382 GATTAGCAAGCGAGGTATGTAGCGGTGCTAACAAGTCTTGAAGTGTAGGCTTAATACTA 1323
| | | | |
QY 2341 CCGCTACACTAGAAAGGACAGATTTTGTGATCTGCGCTTGTGCTGAAGCTTAATCTTCCG 2400
| | | | |
Db 1322 CCGCTACACTAGAAAGGACAGATTTTGTGATCTGCGCTTGTGCTGAAGCTTAATCTTCCG 1263
| | | | |
QY 2401 AAAAAGTTGTAGCTCTTGTATCCGCAAAACAACGCGTGTAGCGGTGTTT 2460
| | | | |

Db 1262 AAAAGAGTTGGTATCTTGATCCGGCAACAAACGACCGCTGGTAGCGGTGTTT 1203
Qy 2461 TGTTCAGAGAGAGATTTAGCGGCAAGAAAAAGATCTCAAGAAATCTTTGATCTT 2520
Db 1202 TGTTCAGAGAGAGATTTAGCGGCAAGAAAAAGATCTCAAGAAATCTTTGATCTT 1143
Qy 2521 TTTCAGGGGCTGACGCTGAGTGAACGAAACCTCACGTTAAGGATTTTGTGATGAG 2580
Db 1142 TTTCAGGGGCTGACGCTGAGTGAACGAAACCTCACGTTAAGGATTTTGTGATGAG 1083
Qy 2581 ATTATCAAAAAGATCTTCACCTAGATCTTTTAAATTTAAATGAATTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGATCTTCACCTAGATCTTTTAAATTTAAATGAATTTTAAATCAAT 1023
Qy 2641 CTTAAGTATATATAGTAAATCTTGTGTCGACAGTTACCAATGCTTATCACTGAGCAC 2700
Db 1022 CTTAAGTATATATAGTAAATCTTGTGTCGACAGTTACCAATGCTTATCACTGAGCAC 963
Qy 2701 TATCTGACGATCTGTATATTCGTTATCCATATGTTGCTGACCTCCCGCTGCTGAT 2760
Db 962 TATCTGACGATCTGTATATTCGTTATCCATATGTTGCTGACCTCCCGCTGCTGAT 903
Qy 2761 AACTACGATACGGAGGGCTTACCATCTGCGCCGACGCTGCAATGATACCGGAGACC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGCGCCGACGCTGCAATGATACCGGAGACC 843
Qy 2821 ACGCTCACCGGCTCAGATTTTATCAGCAATTAACGACGACCGGAGGGCCGAGCGCAG 2880
Db 842 ACGCTCACCGGCTCAGATTTTATCAGCAATTAACGACGACCGGAGGGCCGAGCGCAG 783
Qy 2881 AAGTGTCTGCACTTTATCGGCTCATCAGTCTATTTAATTTGTCGGGAGGCTAG 2940
Db 782 AAGTGTCTGCACTTTATCGGCTCATCAGTCTATTTAATTTGTCGGGAGGCTAG 723
Qy 2941 AGTAAGTATGTCGCACTTATAGTTTGCAGACGTTGTCGCAATGCTACAGGATCT 3000
Db 722 AGTAAGTATGTCGCACTTATAGTTTGCAGACGTTGTCGCAATGCTACAGGATCT 663
Qy 3001 GGTGTCAACGCTCGTGTGTTGATGCTTCACTGACGCTCCGTTCCCAACGATCAAGCG 3060
Db 662 GGTGTCAACGCTCGTGTGTTGATGCTTCACTGACGCTCCGTTCCCAACGATCAAGCG 603
Qy 3061 AGTACATGATCCCGCATGTTGTCGCAAAAAGCGGTTAGCTTCGCTCCGATCGT 3120
Db 602 AGTACATGATCCCGCATGTTGTCGCAAAAAGCGGTTAGCTTCGCTCCGATCGT 543
Qy 3121 TGTGAGAGTATGTCGCGCAGAGTTATCACTGATGTTATGCGACGATGATTAATTC 3180
Db 542 TGTGAGAGTATGTCGCGCAGAGTTATCACTGATGTTATGCGACGATGATTAATTC 483
Qy 3181 TCTTACTGTATGTCATCGGATCGGATGCTTTTCTGTGATGCTGATGATCAACCAAGTC 3240
Db 482 TCTTACTGTATGTCATCGGATCGGATGCTTTTCTGTGATGCTGATGATCAACCAAGTC 423
Qy 3241 ATTGTGAGATGATGTTATGCGGCGACGAGTTGCTTCGCGCGCTCAATACGGGATTA 3300
Db 422 ATTGTGAGATGATGTTATGCGGCGACGAGTTGCTTCGCGCGCTCAATACGGGATTA 363
Qy 3301 TACGCGCGCATATGACAGAACTTTAAAGTCTCATCATTTGGAAGAACTTTCTTGGGGCG 3360
Db 362 TACGCGCGCATATGACAGAACTTTAAAGTCTCATCATTTGGAAGAACTTTCTTGGGGCG 303
Qy 3361 AAAAATCTCAAGATCTTAACTGCTGTTGAGATCAAGTTGATGATCCCACTGTCGAC 3420
Db 302 AAAAATCTCAAGATCTTAACTGCTGTTGAGATCAAGTTGATGATCCCACTGTCGAC 243
Qy 3421 CAATGATCTTACGATCTTTTACTTCAACGAGGTTTCTGGGAGAGAAAGAGAG 3480
Db 242 CAATGATCTTACGATCTTTTACTTCAACGAGGTTTCTGGGAGAGAAAGAGAG 183
Qy 3481 GCAAAATGCGGCAAAAAGGAAATAGGCGCAGACGGAATGTTGAATCTCATACTCTT 3540
Db 182 GCAAAATGCGGCAAAAAGGAAATAGGCGCAGACGGAATGTTGAATCTCATACTCTT 123

Qy 3541 CCTTTTCAATATATGAGATTTATACAGGTTATTTGTCATGAGCGGATACATATT 3600
Db 122 CCTTTTCAATATATGAGATTTATACAGGTTATTTGTCATGAGCGGATACATATT 63
Qy 3601 TGAATGATTTTAAAGAAATTAACAAATAGGAGGTTCCGCGACATTTCCCGAAAGTCC 3660
Db 62 TGAATGATTTTAAAGAAATTAACAAATAGGAGGTTCCGCGACATTTCCCGAAAGTCC 3
Qy 3661 AC 3662
Db 2 AC 1
RESULT 2
AX592967/c AX592967 3637 bp DNA linear PAT 27-JAN-2003
LOCUS Sequence 3 from Patent WO02079468.
DEFINITION AX592967
ACCESSION AX592967
VERSION AX592967.1 GI:27950811
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1 Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
AUTHORS A method of increasing complementarity in a heteroduplex
TITLE Patent: WO 02079468-A 3 10-OCT-2002;
JOURNAL Large Scale Biology Corporation (US)
FEATURES
LOCATION/Qualifiers
SOURCE 1..3637
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 72.5%; Score 2656.4; DB 6; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
Qy 1 CTTAAATGTAAGGCTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTTAAATGTAAGGCTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGGCGGCAAAATCGGCAAAATCCCTTATTAATCAAAAGATGACCGA 120
Db 3577 ATTTTAAACCAATAGGCGGCAAAATCGGCAAAATCCCTTATTAATCAAAAGATGACCGA 3518
Qy 121 GATAGGTTAGAGTGTCTTCCAGTTTGGAAACAGAGTCCATTTAAAGAACTGAGACTC 180
Db 3517 GATAGGTTAGAGTGTCTTCCAGTTTGGAAACAGAGTCCATTTAAAGAACTGAGACTC 3458
Qy 181 CAACGTCAAGGGGCAAAAACCGTCTATCAGGGGATGAGCCACTAGTGAACCATGACC 240
Db 3457 CAACGTCAAGGGGCAAAAACCGTCTATCAGGGGATGAGCCACTAGTGAACCATGACC 3338
Qy 3397 CTAATCAAGTTTATTTGGGGTTCGAGGTGCTGTAAGCACTAATTCGGAACCTTAAAGGAG 3338
Db 241 CTAATCAAGTTTATTTGGGGTTCGAGGTGCTGTAAGCACTAATTCGGAACCTTAAAGGAG 300
Qy 301 CCCCCGATTTTGAAGCTTGAACGGGGAAGCCGGCGGAACGTGGCGAGAAAGGAGGAGAA 360
Db 3337 CCCCCGATTTTGAAGCTTGAACGGGGAAGCCGGCGGAACGTGGCGAGAAAGGAGGAGAA 3278
Qy 361 AGCGAAAGAGAGGGGCGCTAGGGGCTGCGCAAGTGTAGCGGTCAACGCTGCGCTAACAC 420
Db 3277 AGCGAAAGAGAGGGGCGCTAGGGGCTGCGCAAGTGTAGCGGTCAACGCTGCGCTAACAC 3218
Qy 421 CACACCGCGCGCTTAATGCGCGCTTACAGGGCGCGCTCCCATTCGCTACAGGCTGCG 480
Db 3217 CACACCGCGCGCTTAATGCGCGCTTACAGGGCGCGCTCCCATTCGCTACAGGCTGCG 3158
Qy 481 CAACTGTGGGAAGGGGATCGGTCGGGCGCTTCTGCTATTACGCCACGCTGGGAGAG 540

Db 3157 CAACCTGGGAAGGCGATCGGTGGGCGCTCTTGCTAATTACGCACTGCGCAAGG 3098
QY 541 GGGATGTGCTGCAAGGCGATTAAATTGGGTAACGCAAGGTTTCCCACTCAAGCGTTG 600
Db 3097 GGGATGTGCTGCAAGGCGATTAAATTGGGTAACGCAAGGTTTCCCACTCAAGCGTTG 3038
QY 601 TAAAGACGCGCAGGAGCGCGCCCTGTTCAATTACGTTTGAACCCGTGAGAGACG 660
Db 3037 TAAAGACGCGCAGGAGCGCGCGGTAATCACTCATATAGGCGAATTGAGCTCCA 2978
QY 661 GCAGACTCGCGGTGCAATGTGTTTACAGCGTGAATGAGAGATGAAGATCTGAGAC 720
Db 2977 CCGCGGTGCGCGCGCTGAAACTAGTGAATCCCCCGGCTGCGAATTTCTTAATTTGT 2918
QY 721 GCTGCAAGACAGCGACTGATTAAACCTGAAGAATATATATATTGACGTAAGTTAA 780
Db 2917 ATAGTTCAATCCATGTCATGTGTAAATCCAGACGCTGTTACAAACTCAAGAAAGACATGT 2858
QY 781 AGATATATC--ATGCGTAAATTTGACGCAATGGGATGTGTAATACGACTCATATAGGCGCA 838
Db 2857 GGTCTCTGTTTCGTGTGGGATCTTCAAGAGGCAATGTGTGACAGGTAAATGTTGT 2798
QY 839 ATTGGGTACCGGCCCCCTTCGAGTCAAGGTATCGATTAAGCTTGATATCGAATTCCT 898
Db 2797 CTGTAAAGACAGAGGCGCATCGCAATTGGAGTATTTGTGATTAATGCTCTGCTAGTT 2738
QY 899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGCGCCGACCGCGGTGAGCTCCGCTTTT 958
Db 2737 GAACCTTCCACTTTCAATGTTGTGTAAATTTGAAGTTAACTTGAATTCATCTTTT 2678
QY 959 GTTCCCTTAAATGAGGGTTAAATTAGATCCATGCGTCAATTTTACGCACTATCTTCT 1018
Db 2677 GTTGTCTGCGCATGATGATATCATTTGTGTGA- GTTAATAGTTGTATTCAAATTTGTGTCA 2619
QY 1019 AGGGTATCTAGCTGATCAAGATCATATGTCGGGCTTTTTCGGGCTCAAGTCG 1078
Db 2618 AGAATGTTCCATCTTCTTAAATAATCAATCCTTTAACTGATTC-----TATTAACA 2564
QY 1079 CCCAAGCTGCGCTATCTGGGATCGGGAGAGAAAGACCCGTCCTTTCCCGCGAGGT 1138
Db 2563 GGGATATCACTTCAAACTTGACTTC-----AGCAGGTCTTGTAGTTCCGCT 2515
QY 1139 TGAAGCGGATGAAAGATTTCGAGATGACTGCTGCTGCAATTGACCTTGAGCGAA 1198
Db 2514 ATCTTAAATAATTAATTTCTTCTGTATCATACCTTGGGATGGCACTCTTGA AAA 2455
QY 1199 ACGCACGTTTACATGATGATTTGGGAAAGTGTGGCATGACGCTTTAAAGGTAACT 1258
Db 2454 GTATGCGCTTCAATGATCTGGGTATCTTGAAGACATTTGAACCAATMAAGAAAGT 2395
QY 1259 GTTGTTCAGGCCACTGGGATACAGTTGTCGCGGCTTTTCGGGACAAGTTCGGAT 1318
Db 2394 AGTGAACAAGTGTGGCATGGAACAAGTATTTTCCAGTAGTCAATTAATTTAAAGGT 2335
QY 1319 GGTACGCGGAAGCGATCAAGCAACCGAACTACCGGACAGCGGAACTGCGCTGC 1378
Db 2334 AAGTTTTC-----GTAATGTTGCATCACTTCACTCTTCACTGACAGAAATTTGTGC 2280
QY 1379 CGGTGTGCAATTAATGACAGCGGTGCGGCTGGGATATTAAGTACAGGAGAGCGGTA 1438
Db 2279 CCAATTAACATCACTAATCTAATCAAGAAATTTGGAGCACTCCAGTAAAGTTCTTCT 2220
QY 1439 TCTGTGCTGATGCCCGAAGATGACATGATACCCCGTGAATTACCCGCGGCGCGC 1498
Db 2219 CCTTAATCATCGTACCCAGCTTTTGTTCCTTTAGTAGGGTTA-----ATTGCGCGC 2165
QY 1499 TTGGCTAATCATGTCATAGCTTTCTGTGTGAATTTTATTCCTGCTCAATTTCA 1558
Db 2164 TTGGCTAATCATGTCATAGCTTTCTGTGTGAATTTTATTCCTGCTCAATTTCA 2105
QY 1559 CACCAATACAGCGCGGAAGCAATAAGTAAAGCTGGGCTGAATGATGAGCTAA 1618
|||||

Db 2104 CACAACATACAGCGCGAAGCATTAAGTTAAAGCTTGGGCTGCTAATGATGAGCTAA 2045
QY 1619 CTCATATTAATTGGCTTTCGCTCATCTGCGCTTTCAGTCCGGAACCTGTGCTCAG 1678
Db 2044 CTCATATTAATTGGCTTTCGCTCATCTGCGCTTTCAGTCCGGAACCTGTGCTCAG 1985
QY 1679 CTGATTAATGAATTCGCGCAAGCGCGGAGAGGGGTTTGGTATTTGGCGCTCTTCC 1738
Db 1984 CTGATTAATGAATTCGCGCAAGCGCGGAGAGGGGTTTGGTATTTGGCGCTCTTCC 1925
QY 1739 GCTTCTCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1798
Db 1924 GCTTCTCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1865
QY 1799 CACTCAAAAGCGGTAATTAAGGTTATCCACAGATTCAGGGATTAACGAGAAAGACATG 1858
Db 1864 CACTCAAAAGCGGTAATTAAGGTTATCCAGATTCAGGGATTAACGAGAAAGACATG 1805
QY 1859 TGAACAAAGCGCAGCAAAAGCGCAGAACCGTAAAGAGCGCGCTGCTGCTGCTTTC 1918
Db 1804 TGAACAAAGCGCAGCAAAAGCGCAGAACCGTAAAGAGCGCGCTGCTGCTGCTTTC 1745
QY 1919 CATAGCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1978
Db 1744 CATAGCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1685
QY 1979 AACCCGACAGACTAATAAGATACAGAGGCTTCCCGTGAAGCTCCCTGCTGCTGCT 2038
Db 1684 AACCCGACAGACTAATAAGATACAGAGGCTTCCCGTGAAGCTCCCTGCTGCTGCT 1625
QY 2039 CCTGTTCCGACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2098
Db 1624 CCTGTTCCGACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1565
QY 2099 GCGCTTTCATAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158
Db 1564 GCGCTTTCATAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1505
QY 2159 CTGGGCTGTGTCAGAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2218
Db 1504 CTGGGCTGTGTCAGAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445
QY 2219 GCTTGTGCTCAACCGCGTAAAGACGACTTATGCGCACTGAGCAGCACTGTAAAC 2278
Db 1444 GCTTGTGCTCAACCGCGTAAAGACGACTTATGCGCACTGAGCAGCACTGTAAAC 1385
QY 2279 AGAATTAACAGCAGGATATGTAAGCGGTCTACAGATTTTGAAGTGTGCTTAAC 2338
Db 1384 AGAATTAACAGCAGGATATGTAAGCGGTCTACAGATTTTGAAGTGTGCTTAAC 1325
QY 2339 TACGCTTACCTAAGAGACGATTTGTATCTGCGCTGCTGCTGCTGCTGCTGCTGCT 2398
Db 1324 TACGCTTACCTAAGAGACGATTTGTATCTGCGCTGCTGCTGCTGCTGCTGCTGCT 1265
QY 2399 GGAAGAAAGATTTGCTACTCTGTAACCGGCAACAAACCCGCTGCTGCTGCTGCTGCT 2458
Db 1264 GGAAGAAAGATTTGCTACTCTGTAACCGGCAACAAACCCGCTGCTGCTGCTGCTGCT 1205
QY 2459 TTTGTTTGAAGCAGATTAAGCGCAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2518
Db 1204 TTTGTTTGAAGCAGATTAAGCGCAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1145
QY 2519 TTTTCTTACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2578
Db 1144 TTTTCTTACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085
QY 2579 AGATTAACAAGAAAGATCTTCACTGATCTTTTAAATTAAGAAAGAAAGAAAGAAAG 2638
Db 1084 AGATTAACAAGAAAGATCTTCACTGATCTTTTAAATTAAGAAAGAAAGAAAGAAAG 1025
QY 2639 ATCTAAAGTATATATAGTAAAGCTTGTGCTGACAGTTACCAATGCTTAATCATGAGCA 2698
Db 1024 ATCTAAAGTATATATAGTAAAGCTTGTGCTGACAGTTACCAATGCTTAATCATGAGCA 965
|||||

QY 2699 CCTATCTCAGGATCTGTCTATTTGTTGATCCATAGTGGCTGACTCCCGTGTGTAG 2758
 Db 964 CCTATCTCAGGATCTGTCTATTTGTTGATCCATAGTGGCTGACTCCCGTGTGTAG 905
 QY 2759 ATAATCAGATACGGAGGGCTTACCATCTGGCCCAAGTGTGCATGATACCGCGAGAC 2818
 Db 904 ATAATCAGATACGGAGGGCTTACCATCTGGCCCAAGTGTGCATGATACCGCGAGAC 845
 QY 2819 CCACGCTCACCAGCTCCAGATTTATCAGCAATAAACGAGCCGAGAGGGCCGAGCGC 2878
 Db 844 CCACGCTCACCAGCTCCAGATTTATCAGCAATAAACGAGCCGAGAGGGCCGAGCGC 785
 QY 2879 AGAAGTGTCTCTGAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAAGT 2938
 Db 784 AGAAGTGTCTCTGAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAAGT 725
 QY 2939 AGAAGTGTCTCTGAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAAGT 2998
 Db 724 AGAAGTGTCTCTGAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAAGT 665
 QY 2999 GTGGTGTACGCTCGTGGTTGGTATGAGCTTCAATTCAGCTCCGTTCCCAACGATCAAG 3058
 Db 664 GTGGTGTACGCTCGTGGTTGGTATGAGCTTCAATTCAGCTCCGTTCCCAACGATCAAG 605
 QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAGAAAGCGGTAGCTCTTCCGTCTCCGATC 3118
 Db 604 CGAGTTACATGATCCCCCATGTTGTGCAGAAAGCGGTAGCTCTTCCGTCTCCGATC 545
 QY 3119 GTTGTGAGAAATAGTTGGCCGAGTGTATCATCTAGTGTATGGAGAGCTGATAT 3178
 Db 544 GTTGTGAGAAATAGTTGGCCGAGTGTATCATCTAGTGTATGGAGAGCTGATAT 485
 QY 3179 TCTCTTACGTGATGCCATCCGTAAGATGCTTTTCTGTAGTGTGATGATCAACAG 3238
 Db 484 TCTCTTACGTGATGCCATCCGTAAGATGCTTTTCTGTAGTGTGATGATCAACAG 425
 QY 3239 TCATCTGAGAAATAGTTGGCCGAGTGTATCATCTAGTGTATGGAGAGCTGATAT 3298
 Db 424 TCATCTGAGAAATAGTTGGCCGAGTGTATCATCTAGTGTATGGAGAGCTGATAT 365
 QY 3299 AATACCGCGCCATAGAGAACTTTAAAGTGCATCATTTGAGAAAGCTTTCCGGGG 3358
 Db 364 AATACCGCGCCATAGAGAACTTTAAAGTGCATCATTTGAGAAAGCTTTCCGGGG 305
 QY 3359 CGAAACTCTCAAGATCTTACCGCTGTGATCCAGTGTGATGATCAACCTCGTCA 3418
 Db 304 CGAAACTCTCAAGATCTTACCGCTGTGATCCAGTGTGATGATCAACCTCGTCA 245
 QY 3419 CCCAAGTGTATTCAGATCTTTTACCTTCAACGCGTTTCTGGGTAGCAAAAACGGA 3478
 Db 244 CCCAAGTGTATTCAGATCTTTTACCTTCAACGCGTTTCTGGGTAGCAAAAACGGA 185
 QY 3479 AGGCAAAATGCGCAAAAGGAATAAGGGCGACAGGAAATGTTGAATCTCATCTC 3538
 Db 184 AGGCAAAATGCGCAAAAGGAATAAGGGCGACAGGAAATGTTGAATCTCATCTC 125
 QY 3539 TTCTTTTTCATATTAATGAAGATTTATCAGGGTATTTGTCTCATGAGCGATACATA 3598
 Db 124 TTCTTTTTCATATTAATGAAGATTTATCAGGGTATTTGTCTCATGAGCGATACATA 65
 QY 3599 TTTGAATGATTTTGAATAAATGAAGGGTTCGCGCACATTTCCCGGAAAAGTG 3658
 Db 64 TTTGAATGATTTTGAATAAATGAAGGGTTCGCGCACATTTCCCGGAAAAGTG 5
 QY 3659 CCAC 3662
 Db 4 CCAC 1
 RESULT 3
 ASPGREEN1/c
 LOCUS ASPGREEN1 3626 bp DNA linear SYN 15-MAY-1998

DEFINITION Artificial sequences, plasmid vector pGreen-1.
 ACCESSION Y09373
 VERSION Y09373.1 GI:1684624
 KEYWORDS blaTEM gene; lacZ-phoc fusion gene; multiple cloning site; T3 promoter.
 SOURCE ORGANISM synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1 Thaller, M.C., Berlitti, F., Schippa, S., Selan, L. and Rossolini, G.M. Bacterial acid phosphatase gene fusions useful as targets for cloning-dependent insertional inactivation Biotechnol. Prog. 14 (2), 241-247 (1998)
 JOURNAL 98215688
 MEDLINE 9548775
 PUBMED 2 (bases 1 to 3626)
 REFERENCE Rossolini, G.M.
 AUTHORS Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
 TITLE Direct Submission
 JOURNAL Molecular-sez., Microbiologia, Univ. di Siena, via Laterina N.8, 1-53100 Siena, ITALY
 COMMENT Related sequences: X64444 and X52328.
 FEATURES location/Qualifiers
 source 1..3626
 /organism="synthetic construct"
 /mol_type="other DNA"
 /db_xref="taxon:32630"
 /plasmid="pGreen-1"
 129..989
 /gene="blaTEM"
 129..989
 /gene="blaTEM"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA70547.1"
 /db_xref="GI:1684625"
 /translation="MSIQHRYALLIPFAFCIPVPAHPETLVKVAAPDLAGRVY
 IEIDLSNGKILBSFPRERFPMSTFKVLICGAVLSRIDGQEOGLRIHYSONDLYE
 YSPVTEKHLTDGNTVRELCSAAITMSDNTANLLTTTIGAPKEITAFLEHMGDVTFL
 DRWPELNEAIIPNDERPTTPVMAATLRLTGLBELTLASROQLIDMWDADKVAQPL
 LRSLPAGWFPIADKSGAGERSGIIIALPDGKPSRIVIVITTSQATWDERRQIA
 EIGASLIKHW"
 2139..2154
 primer_bind
 /note="pUC M13/Reverse Sequencing primer recognition site"
 2146..2934
 /gene="lacZ-phoc fusion"
 2146..2934
 /gene="lacZ-phoc fusion"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA70548.1"
 /db_xref="GI:1684626"
 /translation="MTWITPRAOULTLTKGNKSSPVAALSLVDPAIPAGNDATTK
 PDLYLVKNEQAIISLKLPPPPVGSIOPLNDQAMYEKGMLENTERGKQAOADADLA
 AGSVATRFSAFEGYPTTEKDSPELYKLLTMMIDAGDLATRSKAEHMRIRPPFYGT
 ETCNTKQKSLSTNGSPSGHTSIGAVTALVLAVNPANDALIEREYQUGSKRVIG
 YHWQSDVARIVISAAVATLHSDPAFOQLAKAKOEFKSKOK"
 2171..2190
 promoter
 /gene="lacZ-phoc fusion"
 /note="T3 promoter"
 ORIGIN
 Query Match 72.4%; Score 2650.6; DB 12; Length 3626;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 3079; Conservative 0; Mismatches 544; Indels 42; Gaps 6;
 QY 1 CTAATTTGAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
 Db 3626 CTAATTTGAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3567
 QY 61 ATTTTAAACCAATAGCGCAAAATCCCTTATTAATCAAAAGAAATGACCGA 120
 Db 3566 ATTTTAAACCAATAGCGCAAAATCCCTTATTAATCAAAAGAAATGACCGA 3507

QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCCTATTTAAAGAACGTGACATC 180
DB 3506 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCCTATTTAAAGAACGTGACATC 3447
QY 181 CAAGCTCAAAAGGCGAAAAACCGCTATTCAGGGCGATGAGCCACTGACGTGAACCAATCAC 240
DB 3446 CAAGCTCAAAAGGCGAAAAACCGCTATTCAGGGCGATGAGCCACTGACGTGAACCAATCAC 3387
QY 241 CTATCAAGTTTGTGGGGGTGAGGTGCGGCTTAAGCACTTAATCGGAACCTTAAAGGAG 300
DB 3386 CTATCAAGTTTGTGGGGGTGAGGTGCGGCTTAAGCACTTAATCGGAACCTTAAAGGAG 3327
QY 301 CCCCCGATTTAAGCTTGAACGCGGAAAGCCGCGAAGCTGCGAGAGAAAGAGAGAA 360
DB 3326 CCCCCGATTTAAGCTTGAACGCGGAAAGCCGCGAAGCTGCGAGAGAAAGAGAGAA 3267
QY 361 AGGAAAGAGAGCGGGGCTAAGAGCGCTGCGCAAGTGAAGCGGTACGCTGCGGTAACAC 420
DB 3266 AGGAAAGAGAGCGGGGCTAAGAGCGCTGCGCAAGTGAAGCGGTACGCTGCGGTAACAC 3207
QY 421 CACACCCGCGCTTAATGCGCCGCTACAGGCGCGTCCATTGCGCAATTCAGGCTGCG 480
DB 3206 CACACCCGCGCTTAATGCGCCGCTACAGGCGCGTCCATTGCGCAATTCAGGCTGCG 3147
QY 481 CAACGTGTGGGAAGGCGGATCGGTGCGGCGCTTTGCTATTAACGCGAGCTGCGGAAAG 540
DB 3146 CAACGTGTGGGAAGGCGGATCGGTGCGGCGCTTTGCTATTAACGCGAGCTGCGGAAAG 3087
QY 541 GGGATGTGTGTCGACGAGCTTAAGTGGTAAACGCAAGGTTTTCCCAAGTACAGACTTG 600
DB 3086 GGGATGTGTGTCGACGAGCTTAAGTGGTAAACGCAAGGTTTTCCCAAGTACAGACTTG 3027
QY 601 TAAACGACGCGCACTGAGCGCGCTGCTTCAATTCAGGTTTTGAACCCGCGAGAGACG 660
DB 3026 TAAACGACGCGCACTGAGCGCGCTGCTTCAATTCAGGTTTTGAACCCGCGAGAGACG 2567
QY 661 GCAGACTCGCGATGCAAAATGTGTTTACAGCGTATAGAGACATGTAAGTCTCGACAC 720
DB 2966 GCAGACTCGCGATGCAAAATGTGTTTACAGCGTATAGAGACATGTAAGTCTCGACAC 2907
QY 721 GCTCGAGAACGCGACGTAGATTAACCTTAGAAAGATATCATATTGTGACGTAACTTAA 780
DB 2906 GCTCGAGAACGCGACGTAGATTAACCTTAGAAAGATATCATATTGTGACGTAACTTAA 2847
QY 781 AGATATCATGCGTAATAATTGACGATGAGTCTGTATTCGACTCACTATTAAGGCGAAT 840
DB 2846 AGATATCATGCGTAATAATTGACGATGAGTCTGTATTCGACTCACTATTAAGGCGAAT 2794
QY 841 TGGGTACCGGAGCGCCCTCTGAGAGTGCACGATTCGATTAAGCTTGATTCGAATTCCTTC 900
DB 2793 TGGGTACCGGAGCGCCCTCTGAGAGTGCACGATTCGATTAAGCTTGATTCGAATTCCTTC 900
QY 901 AGCCCGGAGGATCCACAGTTTAAAGCGCGCGCCACCGCGGTGAGCTTCAGCTTTTGT 960
DB 2733 AGCCCGGAGGATCCACAGTTTAAAGCGCGCGCCACCGCGGTGAGCTTCAGCTTTTGT 2678
QY 961 TCCCTTTAGTAGAGGTTAAATTAGATCCATGCGTCAATTTTACGAGACTAATCTTTAG 1020
DB 2677 TCCCTTTAGTAGAGGTTAAATTAGATCCATGCGTCAATTTTACGAGACTAATCTTTAG 2619
QY 1021 GGTAAATCTAGAGTCACTAGAGATCATATGTCGGGCTTTTTCGGGCTAGTACATCGCC 1080
DB 2618 GGTAAATCTAGAGTCACTAGAGATCATATGTCGGGCTTTTTCGGGCTAGTACATCGCC 2559
QY 1081 CAAGCTGAGGCTATCTGGGATTCGGGAGAGAAAGCCGCGCTTTTCCGCGAGGTTG 1140
DB 2558 CAAGCTGAGGCTATCTGGGATTCGGGAGAGAAAGCCGCGCTTTTCCGCGAGGTTG 2499
QY 1141 AAGCGGCAATGGAAGATTGCGGAGATGACTGCTGCTGATTAACGTTGAGCGAAAC 1200
DB 2498 AAGCGGCAATGGAAGATTGCGGAGATGACTGCTGCTGATTAACGTTGAGCGAAAC 2439

QY 1201 GCAGGTTTACAGATGATTCGGGAAGGTGCGCATGCA---CGCCTTAAACGTGAAC 1257
DB 2438 TCTGCTGTGCTGTGTTTCCGCGCTCGTATATTCGAGATACAGGCTCTTCTCATCAAT 2379
QY 1258 TGTTCGTTACAGCCACCTGGAGATACAGTTTCGTGCGGCTTTTCCGAGACAGTTCCGA 1317
DB 2378 TGTTCGTTACAGCCACCTGGAGATACAGTTTCGTGCGGCTTTTCCGAGACAGTTCCGA 2319
QY 1318 TGGTCAGCCGGAAGCGGATACAGACCCGAAATACCGGCGACAGCCGGAATCGCTG 1377
DB 2318 TGGTCAGCCGGAAGCGGATACAGACCCGAAATACCGGCGACAGCCGGAATCGCTG 2277
QY 1378 CCGGTGTCAGATTAAATGACAGCGGCGCGCGCTGGGATTAATGTCAGCCAGACAGG 1437
DB 2276 CCGGTGTCAGATTAAATGACAGCGGCGCGCGCTGGGATTAATGTCAGCCAGACAGG 2217
QY 1438 ATCTGCTGATATCCCGCAGAAATGACATGATACCCGCTGAGTTAACCGCGGCGCG 1497
DB 2216 ATCTGCTGATATCCCGCAGAAATGACATGATACCCGCTGAGTTAACCGCGGCGCG 2166
QY 1498 CTGCGATATCATGATGATATGATCTGTTTCCGTGGAATTGTTATCCGTCACAAATTC 1557
DB 2165 CTGCGATATCATGATGATATGATCTGTTTCCGTGGAATTGTTATCCGTCACAAATTC 2106
QY 1558 ACACAAATATGAGCGCGAAGCATTAAGTTAAAGCTCGGAGTCTTAATGATGAGCTA 1617
DB 2105 ACACAAATATGAGCGCGAAGCATTAAGTTAAAGCTCGGAGTCTTAATGATGAGCTA 2046
QY 1618 ACTCAATTAATTTGTTGCGCTCATCTGCGCGCTTTTCAAGTCGGGAAACCTGCTGCA 1677
DB 2045 ACTCAATTAATTTGTTGCGCTCATCTGCGCGCTTTTCAAGTCGGGAAACCTGCTGCA 1986
QY 1678 GCTGATTAATGATTCGCGCAACGCGCGGAGAGCGGTTTGGTATGAGGCGCTTTC 1737
DB 1985 GCTGATTAATGATTCGCGCAACGCGCGGAGAGCGGTTTGGTATGAGGCGCTTTC 1926
QY 1738 CGCTTCCTGCTCATCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCGGAGCGGATCAGC 1797
DB 1925 CGCTTCCTGCTCATCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCGGAGCGGATCAGC 1866
QY 1798 TCACTCAAGCGGTTAATGAGTTATTCACAGATCAGGAGATTAAGCGGAAAGAACT 1857
DB 1865 TCACTCAAGCGGTTAATGAGTTATTCACAGATCAGGAGATTAAGCGGAAAGAACT 1806
QY 1858 GTGAGCAAAAGCGCGCAAAAGCGCAGAAACCTTAAGGCGCGTTCGCTGCTGCTTTC 1917
DB 1805 GTGAGCAAAAGCGCGCGCAAAAGCGCAGAAACCTTAAGGCGCGTTCGCTGCTGCTTTC 1746
QY 1918 CCATAGGCTCGCGCCCTGACAGCATCAAAAAATGACAGCTCAAGTCAAGAGTGGCG 1977
DB 1745 CCATAGGCTCGCGCCCTGACAGCATCAAAAAATGACAGCTCAAGTCAAGAGTGGCG 1686
QY 1978 AAACCGGACGAGCTTAAGATCAAGGCGTTTCCCTGGAAGTCCCTGTCGCTC 2037
DB 1685 AAACCGGACGAGCTTAAGATCAAGGCGTTTCCCTGGAAGTCCCTGTCGCTC 1626
QY 2038 TCTGTTGCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGGCT 2097
DB 1625 TCTGTTGCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGGCT 1566
QY 2098 GGGCTTTTCTATAGTCAAGCTGATGATTCAGTTTCGATGAGTCTGCTGCTCA 2157
DB 1565 GGGCTTTTCTATAGTCAAGCTGATGATTCAGTTTCGATGAGTCTGCTGCTCA 1506
QY 2158 GCTGAGCTGTGCAAGAACCCCGGTTCAAGCCGACCGCTGCGCTTATCCGTAATCA 2217
DB 1505 GCTGAGCTGTGCAAGAACCCCGGTTCAAGCCGACCGCTGCGCTTATCCGTAATCA 1446
QY 2218 TCGTCTTAATGATCAACCCGATTAAGACAGCTTAATGCGCATGCGAGAGCACTGATA 2277
DB 1445 TCGTCTTAATGATCAACCCGATTAAGACAGCTTAATGCGCATGCGAGAGCACTGATA 1386
QY 2278 CAGGATTAAGCAGAGAGTATGTAAGCGGTGCTACAGAGTTCTTGAAGTGTGCTTAA 2337

Db 1385 CAGATTAGCAGAGGAGGTATGTAGGGGCTGCTACAGATTTCTTGAAGTGTGGCTTAA 1326
 Qy 2338 CTACGGCTACCTAGAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTT 2397
 Db 1325 CTACGGCTACCTAGAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTT 1266
 Qy 2398 CGGAAAAGATTTGTAGCTCTTGATTCGGGCAAAACAACCGCTGGTGAAGGGTGTGTT 2457
 Db 1285 CGGAAAAGATTTGTAGCTCTTGATTCGGGCAAAACAACCGCTGGTGAAGGGTGTGTT 1206
 Qy 2458 TTTTGTTCAGCAGCAGATTTACGCGAGAAAAAAGATCTCAAGAAGATCTTTGAT 2517
 Db 1205 TTTTGTTCAGCAGCAGATTTACGCGAGAAAAAAGATCTCAAGAAGATCTTTGAT 1146
 Qy 2518 CTTTTCAGGGGCTCTGACGCTCAGTGAACGAAAACTCAGCTTAAGGATTTTGGTAT 2577
 Db 1145 CTTTTCAGGGGCTCTGACGCTCAGTGAACGAAAACTCAGCTTAAGGATTTTGGTAT 1086
 Qy 2578 GAGATTATCAAAAAAGATCTTCACTAGATCTTTTAATTTAAATGAAGTTTAAATC 2637
 Db 1085 GAGATTATCAAAAAAGATCTTCACTAGATCTTTTAATTTAAATGAAGTTTAAATC 1026
 Qy 2638 AATCTAAAGTATATATGAGTAACTTGGCTGACAGTACCAATGCTTAATCAGTGAGAC 2697
 Db 1025 AATCTAAAGTATATATGAGTAACTTGGCTGACAGTACCAATGCTTAATCAGTGAGAC 966
 Qy 2698 ACCATCTCAGCGATCTGTCTATTTCTGTTATCCATAGTGGCTGACTCCCGCTGTGTA 2757
 Db 965 ACCATCTCAGCGATCTGTCTATTTCTGTTATCCATAGTGGCTGACTCCCGCTGTGTA 906
 Qy 2758 GATTAATACGATACGGGAGGGCTTACCATCTGCGCCCGCTGTGCAATGATACCGCGGA 2817
 Db 905 GATTAATACGATACGGGAGGGCTTACCATCTGCGCCCGCTGTGCAATGATACCGCGGA 846
 Qy 2818 CCCAGCTCAGCGGCTCCAGATTTATCAGCAATAAACAGCAGCGGAGGGCCGAGACG 2877
 Db 845 CCCAGCTCAGCGGCTCCAGATTTATCAGCAATAAACAGCAGCGGAGGGCCGAGACG 786
 Qy 2878 CAGAGTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTTATTTGTTCCGGGAGC 2937
 Db 785 CAGAGTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTTATTTGTTCCGGGAGC 726
 Qy 2938 TAGAGTAAAGTGTTCGCCAGTTAATAGTTTGGCGCAAGTGTGGCATTTGCTACAGGCAT 2997
 Db 725 TAGAGTAAAGTGTTCGCCAGTTAATAGTTTGGCGCAAGTGTGGCATTTGCTACAGGCAT 666
 Qy 2998 CGTGGTGCAGCGCTCGTGTGGTATGGCTTCAATCAGTCCGGTTCCCAACGATCAAG 3057
 Db 665 CGTGGTGCAGCGCTCGTGTGGTATGGCTTCAATCAGTCCGGTTCCCAACGATCAAG 606
 Qy 3058 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCAT 3117
 Db 605 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCCGTCCTCCAT 546
 Qy 3118 CGTGTTCAGAAAGTATGGCCGCAAGTGTATCACTCATGTTATGGCAGCATGTCATTA 3177
 Db 545 CGTGTTCAGAAAGTATGGCCGCAAGTGTATCACTCATGTTATGGCAGCATGTCATTA 486
 Qy 3178 TTCTCTTACTGTCAATGCCATCCGTAAATGCTTTTCTGTACTGGTGAAGTCTCAACCA 3237
 Db 485 TTCTCTTACTGTCAATGCCATCCGTAAATGCTTTTCTGTACTGGTGAAGTCTCAACCA 426
 Qy 3238 GTCAATTCGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGGTCAATACGGGA 3297
 Db 425 GTCAATTCGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGGTCAATACGGGA 366
 Qy 3298 TAAATCCGCGCAATACGCAAACTTTAAAGTGTCTCATCTTGAAGAAAGTTCTTCCGG 3357
 Db 365 TAAATCCGCGCAATACGCAAACTTTAAAGTGTCTCATCTTGAAGAAAGTTCTTCCGG 306
 Qy 3358 GCGAAAACCTCTCAAGATCTTACCGGCTGTGAGATCCAGTTCGATGTAAACCATCCGGC 3417

Db 305 GCGAAAACCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCATCTGTC 246
 Qy 3418 ACCCACTGATCTTCAGATCTTTTACTTTTACCAGCGCTTTCTGCGTGAAGAAAAACAG 3477
 Db 245 ACCCACTGATCTTCAGATCTTTTACTTTTACCAGCGCTTTCTGCGTGAAGAAAAACAG 186
 Qy 3478 AAGCGAAAATGCCCAAAAAAGGAATTAAGGCGCACACGGAATTTGAATTCATATCT 3537
 Db 185 AAGCGAAAATGCCCAAAAAAGGAATTAAGGCGCACACGGAATTTGAATTCATATCT 126
 Qy 3538 CTTCCTTTTCAATATTTATTAAGCATTTATCAGGGTATTTGTCTCATGACGATATAC 3597
 Db 125 CTTCCTTTTCAATATTTATTAAGCATTTATCAGGGTATTTGTCTCATGACGATATAC 66
 Qy 3598 ATTGAATGTATTTAGAAAAATTAACAATTAAGGGTGTTCGGCGCATTTCCCGAAAAGT 3657
 Db 65 ATTGAATGTATTTAGAAAAATTAACAATTAAGGGTGTTCGGCGCATTTCCCGAAAAGT 6
 Qy 3658 GCCAC 3662
 Db 5 GCCAC 1

RESULT 4
 LOCUS XKU43957 3624 bp DNA linear SYN 28-JAN-1999
 DEFINITION Expression vector pBI07-MRz, complete cds.
 ACCESSION U43957
 VERSION U43957.1 GI:4097019
 KEYWORDS
 SOURCE Expression vector pBI07-MRz
 ORGANISM Expression vector pBI07-MRz
 REFERENCE Ruiz,J., Wu,C.H., Ito,Y. and Wu,G.Y.
 1 (bases 1 to 3624)
 AUTHORS Ruiz,J., Wu,C.H., Ito,Y. and Wu,G.Y.
 TITLE Direct Submission
 JOURNAL Submitted (27-DEC-1995) Juan Ruiz, Division of
 Gastroenterology-Hepatology, University of Connecticut Health
 Center, Farmington, CT 06030, USA
 COMMENT In vivo expression vector for the synthesis of a ribozyme against
 the core region of the Hepatitis B virus pregenomic RNA. It does
 not recognize the viral core protein but the core region of the
 viral RNA. pBI07-MRz contains a tandem of ten units of the
 ribozyme, separated between them by the target sequence recognized
 by the ribozyme, cloned at the XbaI-HindIII site of pBluescript SK
 (+).

FEATURES
 source
 1..3624 location/Qualifiers
 /organism="Expression vector pBI07-MRz"
 /mol_type="genomic DNA"
 /db_xref="taxon:45892"
 complement(657..1425)
 /note="multiple cloning site"
 complement(721..766)
 /product="hammerhead ribozyme"
 complement(791..836)
 /product="hammerhead ribozyme"
 complement(861..906)
 /product="hammerhead ribozyme"
 complement(931..976)
 /product="hammerhead ribozyme"
 complement(1001..1046)
 /product="hammerhead ribozyme"
 complement(1071..1116)
 /product="hammerhead ribozyme"
 complement(1141..1186)
 /product="hammerhead ribozyme"
 complement(1211..1256)
 /product="hammerhead ribozyme"
 complement(1281..1326)
 /product="hammerhead ribozyme"
 complement(1351..1396)
 /product="hammerhead ribozyme"

Promoter complement (1482..1604)
/note="Lac-Z"
gene complement (2639..3499)
/gene="ampR"
CDS complement (2639..3499)
/gene="ampR"
/function="ampicillin resistance marker"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="A010137.1"
/db_xref="GI:4097020"
/translation="MSIQHRYVALIPFPAPFCIPVAPHEITLVKXDAEDQGA
VELDLSGKILSFPEERPMSTFKVLCAVLSRIDAGROGRTHISQNDLVE
YSPVTEKHLDGWTVEELCSAATMSDNTAANLLITIGPEKELTFLHNGDHVTRL
DRPELNEAIPIVDERDITTPVAMATTKLKLIGELTLASQQLIDWEDAKVAGPL
LRSAIPAGFIADKSGAERGRGIIIALGPDGKPSRIVIVITTSQATMBERNQIA
EIGASLIXIM"

ORIGIN

Query Match 72.2%; Score 2644.8; DB 12; Length 3624;
Best Local Similarity 84.3%; Pred. No. 0;

Matches 3096; Conservative 0; Mismatches 512; Indels 64; Gaps 8;

QY 1 CTAAATTGTAAGCGTTAAATTTGTTAAATTCGCGTTAAATTTGTTAAATCAGCTC 60
DB 4 CTAAATTGTAAGCGTTAAATTTGTTAAATTCGCGTTAAATTTGTTAAATCAGCTC 63
QY 61 ATTTTAAACAATAGCGCGAATCGGCAAAATCCCTTAATTAATTAAGATGACCGA 120
DB 64 ATTTTAAACAATAGCGCGAATCGGCAAAATCCCTTAATTAATTAAGATGACCGA 123
QY 121 GATAGGTTAGTGTGTTTCCAGTTTGAACAAGATCCACTATTAAAGACGTGAC 180
DB 124 GATAGGTTAGTGTGTTTCCAGTTTGAACAAGATCCACTATTAAAGACGTGAC 183
QY 181 CAACGTCAAGGCGCAAAACCGTATCAGGGGAGATGCCCACTACGTGAACCATACC 240
DB 184 CAACGTCAAGGCGCAAAACCGTATCAGGGGAGATGCCCACTACGTGAACCATACC 243
QY 241 CTAATCAAGTTTTTGGGGTGCAGGTGCGGTAAAGCACTAAATCGAACCCTAAAGGAG 300
DB 244 CTAATCAAGTTTTTGGGGTGCAGGTGCGGTAAAGCACTAAATCGAACCCTAAAGGAG 303
QY 301 CCCCAGATTAGAGCTTGAACGGGGAAGCCGGCGCAAGTGGCGAAGAAAGGAGAA 360
DB 304 CCCCAGATTAGAGCTTGAACGGGGAAGCCGGCGCAAGTGGCGAAGAAAGGAGAA 363
QY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGAGCAAGTGAAGCGGTACAGCTGCGCTAAC 420
DB 364 AGCGAAAGAGCGGGCGCTAGGGCGCTGAGCAAGTGAAGCGGTACAGCTGCGCTAAC 423
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTTGCGCATCAGGTG 480
DB 424 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTTGCGCATCAGGTG 483
QY 481 CAACGTGTGGAAAGGGGATCGGTGCGGGCTCTTGCCTAATTAACGCAACCTGCGAAAG 540
DB 484 CAACGTGTGGAAAGGGGATCGGTGCGGGCTCTTGCCTAATTAACGCAACCTGCGAAAG 543
QY 541 GGGATGCTGCAAGGCGATTAAAGTTGGGTAACGCCAGGGTTTTCCAGTCAAGAGCTTG 600
DB 544 GGGATGCTGCAAGGCGATTAAAGTTGGGTAACGCCAGGGTTTTCCAGTCAAGAGCTTG 603
QY 601 TAAACGACGCGCAGT-----AGCGCGCTCGTTCAATTCAGTTTGAACCCGTGGA 654
DB 604 TAAACGACGCGCAGTGTGTAATGACATCACTAATAGGCGAATGGATGCGGCGCC 663
QY 655 GGAAGGAGAGCTCGCGGTGCAATGTGTTTACAGCGTGAATGAGAGATGAAGATGCT 714
DB 664 CCCCCTCGAGGTGCAAGGTATGATAGCTTGAAGTTCAAGAGAGATCTGAGATCCAG 723
QY 715 CGACACGCTGCAAGACGCACTGATTAACCTGAAAGATATCATATTGTGACGTA 774

DB 724 CTTAAGATTTCGCTCAGGAGCTCATGATCTCT--GAGCATTTGCTTAGTTCAGAGAG 780
QY 775 GGTAAAGATATATCATCGTAAATTTGACCATGGGATCTGTATATACACTCATTAAG 834
DB 781 CTCTGATTCAGAGCTTAGAGTTTCGCTCCTCAGGACTCATGCTCTTAGCATTTGCTTA 840
QY 835 GCGAATTTGGATACGGGCGCCCGCTGAGGTGACGGTATTCGATTAAGTTGATATGAAT 894
DB 841 GTTCAGAGACTCTGATTCAGAGCTTAGAGTTTCGCTCCTCAGGAG----- 886
QY 895 TCTCGACAGCCCGGGGAGTCCACTAGTTTGAAGCGGCGCCCAACCGGGGTGAGCTCCAGC 954
DB 887 CTATCAGTCTTAGAGATTTGCTAGTTTCAGGAGA-----C 921
QY 955 TTTTGTTCCTTTAGTAGAGGTTAATTTAGATCCATGCGTCAATTTTACGACGATCTCT 1014
DB 922 TCTGATTCAGAGCTTAGAGTTTCGCTCAGGACTCATGCTCTAGCATTTGCTTAG 981
QY 1015 TTCTAGAGTTAATCTAGCTGATTCAGAGATCATATGTCGGCTTTTTCGGCTCAGTC 1074
DB 982 TTCAGAG-----ACTGGAATTCAGGCTTAGAGTTTCTC--CTCAGGACTCATCA 1033
QY 1075 ATGCCCAAGCTGGCGCTAATCTGGGCATCGGGAGAGAAAGCCCGTCTTTCCCGCG 1134
DB 1034 GTCTGAGCATTTGCTAATTTAGAGACTCTGATCCAGGCTTAGAGTTTCTCTCAG 1093
QY 1135 AGGTGAAGCGGATGAGAAAGATTGCGCAGAGATGACTGTGCTGATTAAGCTTGAAGC 1194
DB 1094 GGAATCATATGCTCTAGAGATTTGCTAATTTAGAGAGACT--CTGATTCAGGCTTAGAGT 1152
QY 1195 GAAAAGCAGCTTACATGATGATTCGGGAAGGTGTGGCCATGACAGCCCTTAAACGCTG 1254
DB 1153 TTGCTCTCAGGACTCATGACTAGTCTGATGAGATTTCTAGTTCAAGAGACTTGGATCCA 1212
QY 1255 AACTTTGTTGAGGCCACTGGGATACAGTTGCTGCGGCTTTCCGACACAGTTTC 1314
DB 1213 GCTTAAGATTTGCTCTCAGGACTCATGACTAGTCTGATGAGATTTCTAGTTCAAGAGACT 1272
QY 1315 GGAATGTCAGCCGAGGAGCTATCAGAACCGGAACATTAACCGGAGACAGCCGAACTGCC 1374
DB 1273 CTGATTCAGAGCTTAGAGTTTGTCTCTCAGGACTCATGATCTGAGGATTTGCTAAGT 1332
QY 1375 GTGCCGTGTCAATTAATGAACAGCGGTGCGGCTGGGATTTACGTACAGAGAGAG 1434
DB 1333 TCAGAGAGCTGTGATCAAGCTTTAAGTTTGTCTTCAAGGACTCATGAGCTCGAGAG 1392
QY 1435 GGTATCTGCTGATCCGCAAGAAATGACATGAGTACCCCGTGAATTAACCGGCGG- 1493
DB 1393 ATTGTGAAGCGGCGCCGACCGCGGTGAGCTTCAGCTTTGTTCTTTAGTGAAGGT 1452
QY 1494 -----GCGCTTGGCGTAAATCAATGATGCTGTTCTGTGTAATTTGATCCGC 1547
DB 1453 TAATTTGACACTTGGGTATCAATGATGCTGATGCTGTTCTGTGTAATTTGATCCGC 1512
QY 1548 TCACAAATTCACACAAATACAGCGCGGAGAGTAAAGTTAAAGCTTGGGGTCTTAAT 1607
DB 1513 TCACAAATTCACACAAATACAGCGCGGAGAGTAAAGTTAAAGCTTGGGGTCTTAAT 1572
QY 1608 GAGTGAAGCTTACATTAATTTGGTTGCGTCACTGCGGCTTTCCAGTGGGAAAC 1667
DB 1573 GAGTGAAGCTTACATTAATTTGGTTGCGTCACTGCGGCTTTCCAGTGGGAAAC 1632
QY 1668 TGTGCGGACAGCTGATTAATTAATTCGCAACCGCGGGAGAGAGCGGTTTGGTATG 1727
DB 1633 TGTGCGGACAGCTGATTAATTAATTCGCAACCGCGGGAGAGAGCGGTTTGGTATG 1692
QY 1728 GGGGCTTTTCGCTTCTGCTCATCTGACTGCTGCGGTGTTTGGTGGGGGAG 1787
DB 1693 GGGGCTTTTCGCTTCTGCTCATCTGACTGCTGCGGTGTTTGGTGGGGGAG 1752
QY 1788 CGGTATCGCTTACCTCAAGCGGATTAATGAGTTATTCAGAGATCAAGGAGTAAACGAG 1847

Dp	1753	CGGTACAGCTCAC	CTCAAAAGGCGTAATACGGTTATCCACGAATCAGAGGATACGAG	1812
Qy	1848	GAAAGAACATGTGAGCAAAAGGCCACGCAAAAGGCCAGAACCTGTAAAAAGCCCGGTTGC		1907
Dp	1813	GAAAGAACATGTGAGCAAAAGGCCACGCAAAAGGCCAGAACCTGTAAAAAGCCCGGTTGC		1872
Qy	1908	TGGCGTTTTCATATGAGGCTCGGCCCCCGTGCAGAGCATCAGAAAAATCGAGGCTCAAGTC		1967
Dp	1873	TGGCGTTTTCATATGAGGCTCGGCCCCCGTGCAGAGCATCAGAAAAATGAGGCTCAAGTC		1932
Qy	1968	AGAGGTGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCGAAGCTCC		2027
Dp	1933	AGAGGTGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCGAAGCTCC		1992
Qy	2028	TCGTGCGCTCTCTCTGTTCCGACCCCTCGGCTTACCGGATATCTGTCCGCTTTCCTCTT		2087
Dp	1993	TCGTGCGCTCTCTCTGTTCCGACCCCTCGGCTTACCGGATATCTGTCCGCTTTCCTCTT		2052
Qy	2088	CGGGAAGGATGCGCTTCTCATATGAGTCAACCTGTAGAGTATCTCAGTTCCGTTAGTTCG		2147
Dp	2053	CGGGAAGGATGCGCTTCTCATATGAGTCAACCTGTAGAGTATCTCAGTTCCGTTAGTTCG		2112
Qy	2148	TTTCGCTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTCAAGCCGACCGGCTGAT		2207
Dp	2113	TTTCGCTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTCAAGCCGACCGGCTGAT		2172
Qy	2208	CCGGTAACTATCGTCTTGAGTCCAACCCGCGTAAACACAGACTTATCCGCACTGACAG		2267
Dp	2173	CCGGTAACTATCGTCTTGAGTCCAACCCGCGTAAACACAGACTTATCCGCACTGACAG		2232
Qy	2268	CCACTGTGTAAACAGATTAGCAGAGCGAGATATGTATAGGCGGTCTACAGATTCTTGAAGT		2327
Dp	2233	CCACTGTGTAAACAGATTAGCAGAGCGAGATATGTATAGGCGGTCTACAGATTCTTGAAGT		2292
Qy	2328	GGTGGCCCTAACTACGAGCTACACTAGAAAGACAGTATTTGATATCGGCGCTGTGTCGAGC		2387
Dp	2293	GGTGGCCCTAACTACGAGCTACACTAGAAAGACAGTATTTGATATCGGCGCTGTGTCGAGC		2352
Qy	2388	CAGTTTACCTTCGGAAGAAAGAGTTGGTAGCTTGTATCCGCGCAACAAACACCGCTGGTA		2447
Dp	2353	CAGTTTACCTTCGGAAGAAAGAGTTGGTAGCTTGTATCCGCGCAACAAACACCGCTGGTA		2412
Qy	2448	GCGGTGTTTTTTTTGTTTGGCAAGCAGAGTTTAGCGCGCAAAAAAGAGTCTCAGAG		2507
Dp	2413	GCGGTGTTTTTTTTGTTTGGCAAGCAGAGTTTAGCGCGCAAAAAAGAGTCTCAGAG		2472
Qy	2508	ATCCCTTGATCTTTTTCTTACGCGGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTAAGGA		2567
Dp	2473	ATCCCTTGATCTTTTTCTTACGCGGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTAAGGA		2532
Qy	2568	TTTTTGTCATGAGATTATCAAAAGAGATCTTCACTAGATCTTTTAAATTAAAGATGA		2627
Dp	2533	TTTTTGTCATGAGATTATCAAAAGAGATCTTCACTAGATCTTTTAAATTAAAGATGA		2592
Qy	2628	GTTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTTACCAATGCTTAA		2687
Dp	2593	GTTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTTACCAATGCTTAA		2652
Qy	2688	TCAGTGAAGGCACTATCTCAGCGATCTGTCTAATTTCCGTATCCATAGTGGCTGACATCC		2747
Dp	2653	TCAGTGAAGGCACTATCTCAGCGATCTGTCTAATTTCCGTATCCATAGTGGCTGACATCC		2712
Qy	2748	CCGTCTGTGTAGTAACTACGATACGAGGAGGCTTACCATCTGAGCCCAAGTCTCAATGA		2807
Dp	2713	CCGTCTGTGTAGTAACTACGATACGAGGAGGCTTACCATCTGAGCCCAAGTCTCAATGA		2772
Qy	2808	TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACACGACCGGAA		2867
Dp	2773	TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACACGACCGGAA		2832
Qy	2868	GGGCGAGCGGAGAAAGTGCTCTCTCAACTTATACCGCTCCATCTCAGTCTTAAATTGTT		2927
Dp	2833	GGGCGAGCGGAGAAAGTGCTCTCTCAACTTATACCGCTCCATCTCAGTCTTAAATTGTT		2892

QY	2928	GCCGGGAAGCTAAGTAAGTATGCCAGTTAATAGTTTGGCAAGCTTGTGCATTG	2987
Db	2893	GCCGGGAAGCTAAGTAAGTATGCCAGTTAATAGTTTGGCAAGCTTGTGCATTG	2955
QY	2988	CTACAGGCATCGTGGTGTCAAGCTCGTGTGTGGTATGGCTTCAATTACGTCGGATCC	3047
Db	2953	CTACAGGCATCGTGGTGTCAAGCTCGTGTGTGGTATGGCTTCAATTACGTCGGATCC	3012
QY	3048	AACGATCAAGCGCAAGTTACATGATCCCCCATGTTGTGCATAAAAAGCGGTTAGCTCTTCG	3107
Db	3013	AACGATCAAGCGCAAGTTACATGATCCCCCATGTTGTGCATAAAAAGCGGTTAGCTCTTCG	3072
QY	3108	GTCCCTCCGATGGTGTGCAAGTAAGTTGGCGGAGTTATCACTCATGGTTATGGCAG	3167
Db	3073	GTCCCTCCGATGGTGTGCAAGTAAGTTGGCGGAGTTATCACTCATGGTTATGGCAG	3133
QY	3168	CACGTCATAAATCTCTTACTGTCAATGCCATCCGTAAAGATGCTTTTCTGTGACTGGTAGT	3227
Db	3133	CACGTCATAAATCTCTTACTGTCAATGCCATCCGTAAAGATGCTTTTCTGTGACTGGTAGT	3192
QY	3228	ACTCAACCAATCATCTTGAGATATAGTATAGCGGCCAGAGTTGCTCTTGGCCGGCGT	3287
Db	3193	ACTCAACCAATCATCTTGAGATATAGTATAGCGGCCAGAGTTGCTCTTGGCCGGCGT	3252
QY	3288	CAATACGGGATTAATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCTTGGAAAAAC	3347
Db	3253	CAATACGGGATTAATACCGCGCACATAGCAGAACTTTAAAGTGCTCATCTTGGAAAAAC	3312
QY	3348	GTTCTTCGGGGCGAAAACTCTCAAGATCTTACCGCTGTGGAGAATCCAGTTGATGTAAAC	3407
Db	3313	GTTCTTCGGGGCGAAAACTCTCAAGATCTTACCGCTGTGGAGAATCCAGTTGATGTAAAC	3372
QY	3408	CCACTCGTGACCCCACTGATCTTCAGCATTTTTACTTTACCAAGGCTTCTGGGTGAG	3467
Db	3373	CCACTCGTGACCCCACTGATCTTCAGCATTTTTACTTTACCAAGGCTTCTGGGTGAG	3432
QY	3468	CAAAAACAGAAAGCAAAAATGCCGCAAAAAAGGAAATAAAGGCGACACGAAATGTTGAA	3527
Db	3433	CAAAAACAGAAAGCAAAAATGCCGCAAAAAAGGAAATAAAGGCGACACGAAATGTTGAA	3492
QY	3528	TACTCATACTCTTCCTTTTCAATATATTATGAACATTTATCAGGGTTATTGTCTCAAGA	3587
Db	3493	TACTCATACTCTTCCTTTTCAATATATTATGAACATTTATCAGGGTTATTGTCTCAAGA	3552
QY	3588	GCGGATACATATTGGAATGTTTAAAGAAAAATAACCAATAGGGGTTCCCGGCACATTTCC	3647
Db	3553	GCGGATACATATTGGAATGTTTAAAGAAAAATAACCAATAGGGGTTCCCGGCACATTTCC	3612
QY	3648	CCCGAAAAAGTGC 3659	
Db	3613	CCCGAAAAAGTGC 3624	
RESULT 5			
LOCUS	A20702	3681 bp	DNA linear PAT 22-AUG-1994
DEFINITION	PSEB5 DNA sequence.		
ACCESSION	A20702		
VERSION	A20702.1	GI:583404	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 3681).		
AUTHORS			
TITLE	CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS		
JOURNAL	USE AS AN IMMUNOGEN AND VACCINE		
FEATURES	Patent: WO 9113155-A 6 05-SEP-1991;		
	Location/Qualifiers		
	1..3681		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		

```

TITLE          CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
JOURNAL        USE AS AN IMMUNOGEN AND VACCINE
FEATURES       Patent: WO 91/3155-A 6 05-SEP-1991;
SOURCE         Location/Qualifiers
               1. .3681
               /organism="synthetic construct"
               /mol_type="unassigned DNA"

```


ORIGIN /db_xref="taxon:32630"

Query Match 72.1%; Score 2639; DB 6; Length 3681;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

3 AAATGTAGAGCTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCACTCAT 62
1 AAATGTAAAGCTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCACTCAT 60
63 TTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATTAATCAAAAGATAGACCGAGA 122
61 TTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATTAATCAAAAGATAGACCGAGA 120
123 TAGGGTTAGAGCTTTTCCAGTTTGGAACAAAGATCCATATTAAGAAGCTGACCTCA 182
121 TAGGGTTAGAGCTTTTCCAGTTTGGAACAAAGATCCATATTAAGAAGCTGACCTCA 180
183 AGCTCAAAAGGCGAAAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 242
181 AGCTCAAAAGGCGAAAAACCGCTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCT 240
243 AATCAAGTTTTTGGGGTGCAGGTGCCGTAAAGCATTAATCGGAACCTTAAAGGAGCC 302
241 AATCAAGTTTTTGGGGTGCAGGTGCCGTAAAGCATTAATCGGAACCTTAAAGGAGCC 300
303 CCGCATTTAGAGCTTGAACGGGGAAGCGCGGAACGTGGCGAAGAAAGAGGAAAG 362
301 CCGCATTTAGAGCTTGAACGGGGAAGCGCGGAACGTGGCGAAGAAAGAGGAAAG 360
363 CGAAAGAGCGGGCGCTAGAGCGCTGGCAAGTGAAGCGGTCAAGCTCGCGTAAACCA 422
361 CGAAAGAGCGGGCGCTAGAGCGCTGGCAAGTGAAGCGGTCAAGCTCGCGTAAACCA 420
423 CACCGCGCGGCTTAATGCCCGCTTACAGGCGCGTCCCATTTGCCCATCAGGCTGCGCA 482
421 CACCGCGCGGCTTAATGCCCGCTTACAGGCGCGTCCCATTTGCCCATCAGGCTGCGCA 480
483 ACTGTTGGGAAGGCGATCGGTGCGGCGCTTTCGCTAATTAACCGCAAGCTGGCGAAGG 542
481 ACTGTTGGGAAGGCGATCGGTGCGGCGCTTTCGCTAATTAACCGCAAGCTGGCGAAGG 540
543 GATGTGCTGCAAGCGCATTAAGTTGGGTAAAGCGCAAGGTTTTCCAGTCAAGCTTTGA 602
541 GATGTGCTGCAAGCGCATTAAGTTGGGTAAAGCGCAAGGTTTTCCAGTCAAGCTTTGA 600
603 AAACGACGCGCAGTGAAGCGCGCTGCTCAATCACTTTTGAACCCGTGAGAGAGCGGC 662
601 AAACGACGCGCAGTGAAGTTGTAATACGACTCACTAATAGGCGCAATTGAGCTCCACCGCG 660
663 AGACTCGCGGTGCAAAATGTTTTCAGCGTGAAGAGATGAAGATGCTGCACACGC 722
661 GTGGCGGCGCTCTAGATGTGCATCTCAAGTACATCTGCTGATGCCCATATGTTA 720
723 TGCAGAACGCGCAGTGAAGTTAACCTTAAGAAATATATATTTGACAGTACGTTAAAG 782
721 AGCAGATATATACATCCGCTATCGTACGTGAATGAGTCAATGAGTCCCGCACACCC 780
783 AATATCATGCTTAATTAATGACGATGGATCTGTAAATACGATCACTAATAGGCGCAATTG 842
781 GCCAACACCGCGTGAAGCGCGCTGAGAGGCTTGTCTGCTCCGGGATCCGCTTAAGACA 840
843 GGTACGCGGCGCGCGCTGAGAGTGCAGGATTCGA-----TGAAGTGAATTCGAATTTC 896
841 AGCTGTGACCGTCTCGGAGGCTGACGTGTCAGAGGTTTTCACCGTATCAACGAAAGC 900
897 CTGACAGCGCGGAGATCACTAGTTCTAGAGGCGCGCAACCGCGTGAAGCTCCAGCTT 956
901 CCGAGAGCAGTGAAGTGCAGTCTTTGTAGCAAAATTCGCTTAAAGTAAAGCAATTTGCTG 960
957 TTGTTCTCTTAAAGTGAAGGTTAATGATCCCAT-----GCTCAATTTTAAAGCAAGC 1009

961 TAAAGTCGACTGTGCGATCAACCGCTTCCAGTACGACAGAAAGCAATTTGATGTA 1020
1010 TATCTTTCTAGAGTTAATCTAGCTGATCAGGATCATATCTGCGGCTCTTTTCCGCT 1069
1021 TTTGAGAGAAAGATGCGAGAAATCAATATCAATTAAGTGAATCTTTTGTGTGCT 1080
1070 CAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGAGAGAAAGACCGGCTTTTC 1129
1081 TCGACATGGGTATCTCTCATGTTTGAATGGCCCTAAGAGATCCGCAAGCTTGCAATGCC 1140
1130 CCGGAGGTTGAAGCGGCAATGGAAGATTTGCGAAGATGACTGCTGCTGATTCAGCT 1189
1141 TGAAGTCACTCTAAGAGATCCCGACCTCGACCGCATTAATATGTTTTCCGTAAT 1200
1190 TGAAGCAAAACGACGTTTACATGATGATTCGGGAAGGTTGGCCATGACAGCTTTA 1249
1201 TCAAGCGCTTTCATGATGAGACAGGCGCTTT--GAATGTTGACGGGATGAATATTAAG 1258
1250 CCGTGAACCTGTTCTGTTCAAGCCCACTGGGATACAGTTGCTGCGGCTTTTCCGACACA 1309
1259 CAATGACGCAAGCAATTAATCAACAG--AGCAAGAAAGGAGGATATCCACAAAGT 1315
1310 GTTCCGGAATGTCAGCCCGAAGCGATCAGCAACCGGAACAATACCGCGCAGCAG----- 1363
1316 CCAAGCTACCAATTAAGCAAGGCTCAACGACGACGACGACGAGAGCGGTCAATAGCA 1375
1364 CCGAATCGCGCTGCGGCTGTCAGATTAATGACAGCGGTGCGGCTGGGATATTACGT 1423
1376 TCCAAACTTTGTTACTCGTCGACGAAATTCGAATATCTTTCGTTAAATCCAAAGCGCAG 1435
1424 CAGCGAGACGGGTAT---CTGCGTGAATGCGCGCAAAATGACATGATATCCCGCTG 1479
1436 AAGCTGATGAAATTCGACTCGAGGGGGGCGCGTACCAGCTTTGTTCCCTTTA 1495
1480 AGTTACCGGCGGCGGCTTGGCGTATCAATGATGATGATGCTTTCTGTGTGAATTTG 1539
1496 GTGAGGTTAATTCGAGCTTGGCGTATCATGTCATAGCTGTTTCTGTGTGAATTTG 1555
1540 TTAATCGCTCAATTTCAACAACAATACGAGCGGGAAGATTAAGTAAAGCTGAGGG 1599
1556 TTAATCGCTCAATTTCAACAACAATACGAGCGGGAAGATTAAGTAAAGCTGAGGG 1615
1600 TGCCTAATGAGTGAAGCTAATCAATTAATGCTGCTGCTCACTGCGCGCTTTTCAAGTC 1659
1616 TGCCTAATGAGTGAAGCTAATCAATTAATGCTGCTGCTCACTGCGCGCTTTTCAAGTC 1675
1660 GGGAAACCTGTCTGCGCAGCTGCAATTAATGATGAGCGCAACGCGGGGAGAGGCGGTT 1719
1676 GGGAAACCTGTCTGCGCAGCTGCAATTAATGATGAGCGCAACGCGGGGAGAGGCGGTT 1735
1720 GCGTATTGGGCGCTCTTCCGCTTCTGCTCACTGACTGCTGCGCTGCTGCTTCCGCT 1779
1736 GCGTATTGGGCGCTCTTCCGCTTCTGCTCACTGACTGCTGCGCTGCTGCTTCCGCT 1795
1780 GCGGCGAGCGGTATCAGCTCACTCAAAAGGCGTAAATGCTGTTATCCACAAATTAAGGGA 1839
1796 GCGGCGAGCGGTATCAGCTCACTCAAAAGGCGTAAATGCTGTTATCCACAAATTAAGGGA 1855
1840 TAAAGCAGAAAGAAATGTAAGCAAAAGGCGCAAAAGGCGCAGAACGCTTAAGAAAGGC 1899
1856 TAAAGCAGAAAGAAATGTAAGCAAAAGGCGCAAAAGGCGCAGAACGCTTAAGAAAGGC 1915
1900 CCGGTTGCTGGCGTTTTTTCATAGGCTCGCGCGCGCTGACGAGATCACAAAATTCGACG 1959
1916 CCGGTTGCTGGCGTTTTTTCATAGGCTCGCGCGCGCTGACGAGATCACAAAATTCGACG 1975
1960 CTCAAGTCAGAGTGGGAAACCGCAAGAGCTATAAAGTATACAGGCGTTTTCCCTG 2019
1976 CTCAAGTCAGAGTGGGAAACCGCAAGAGCTATAAAGTATACAGGCGTTTTCCCTG 2035
2020 AAGTCTCTGTGCGCTCTCTGTTCCGACCTGCGCTTAACCGGATACCTGTCCGCTT 2079
2036 AAGTCTCTGTGCGCTCTCTGTTCCGACCTGCGCTTAACCGGATACCTGTCCGCTT 2095

QY 183 AAGTCAAGAGGAGGAAAAACCGTCTATCAGGCGGATGCGCCACTACGTAAACCATCACCCT 242
| | | | |
Db 181 AAGTCAAGAGGAGGAAAAACCGTCTATCAGGCGGATGCGCCACTACGTAAACCATCACCCT 240
| | | | |
QY 243 AATCAAGTTTTTTTGGGGGTGAGAGTGCCTGTAAGACATTAATCGGAACCTTAAGAGGAGCC 302
| | | | |
Db 241 AATCAAGTTTTTTTGGGGGTGAGAGTGCCTGTAAGACATTAATCGGAACCTTAAGAGGAGCC 300
| | | | |
QY 303 CCGGATTTAGAGCTTGAAGGAGGAAAGCCGCGCAACGTGCGAGAAAGAAAGAGAGAAAG 362
| | | | |
Db 301 CCGGATTTAGAGCTTGAAGGAGGAAAGCCGCGCAACGTGCGAGAAAGAAAGAGAGAAAG 360
| | | | |
QY 363 CGAAAGAGCGGAGCGCTAGAGGCGCTGAGAGTGAAGCGGTACCGCTGCGCAATTCAGGCTGAGCA 422
| | | | |
Db 361 CGAAAGAGCGGAGCGCTAGAGGCGCTGAGAGTGAAGCGGTACCGCTGCGCAATTCAGGCTGAGCA 420
| | | | |
QY 423 CACCGCGCGCTTAATGAGCGCGCTACAGGCGCGCTCCATTGCGCAATTCAGGCTGAGCA 482
| | | | |
Db 421 CACCGCGCGCTTAATGAGCGCGCTACAGGCGCGCTCCATTGCGCAATTCAGGCTGAGCA 480
| | | | |
QY 483 ACTGTGGGAAGGAGGAGTCCGCTGAGGCGCTCTTGGCTATTAGCGCAAGCTGCGAGAAAGAGG 542
| | | | |
Db 481 ACTGTGGGAAGGAGGAGTCCGCTGAGGCGCTCTTGGCTATTAGCGCAAGCTGCGAGAAAGAGG 540
| | | | |
QY 543 GATGTGCTGCAAGGCGATTAAGTTAGGTTAAGCGCAAGGTTTTCCAGTCAAGAGCTTGA 602
| | | | |
Db 541 GATGTGCTGCAAGGCGATTAAGTTAGGTTAAGCGCAAGGTTTTCCAGTCAAGAGCTTGA 600
| | | | |
QY 603 AAAGGAGCGCGAGTGAAGCGCGCGCTCGTTCAATTCAGTTTTTGAACCCGTGAGAGAGCGAGC 662
| | | | |
Db 601 AAAGGAGCGCGAGTGAAGTGTAAATGATACATCACTATAGGCGAAATTTGAGACTCCACCGCG 660
| | | | |
QY 663 AGACTCGCGGTGCAAAATGTGTTTTACAGCGTGAAGAGCAATGAAGATCTGACAGCG 722
| | | | |
Db 661 GTGCGCGCGCGCTGATGATGTGTGCACTCTAGTAAATCAATTCGTGTGATGCGCATAGTTA 720
| | | | |
QY 723 TGCAGAAACAGCAGCTAGATTAACCTTGAAGAAAGTAAATCATATTGTGACGTAAAG 782
| | | | |
Db 721 AGCAGATATATACCTCCGCTATGCTACGTAAGTGAAGGCTGAGGCTGCGCCCGGACAGCC 780
| | | | |
QY 783 ATATATCATGGGTAAATTTAGAGCAATGGAGATCTGTAAATCACTCATATAGGCGCAATTG 842
| | | | |
Db 781 GCCAATCCCGCTGAGCGCGCTGACGCGGCTGTGTCTCCGCGCATCCGCTTAACAGACA 840
| | | | |
QY 843 GGTACCGGCGCGCGCTCGAGGTGACAGGTATCGA-----TAAAGTGAATATCGAATTC 896
| | | | |
Db 841 AGCTGTGACCGTCTCCGCGAGCTGCAATGTGTCAAGAGTTTTTCAACGCTCATCACCGAAACG 900
| | | | |
QY 897 CTGCAAGCGCGGAGATTCATAGTTCTAGAGCGCGCGCAACGCGGCTGAGAGCTCCAGTT 956
| | | | |
Db 901 CCGAGGCAATTAAGTGTGAGTGTGTGTGAGCAATTCGTCCCTTAAGTAAAGCAATTGCTG 960
| | | | |
QY 957 TTGTTCCCTTTAGTGAAGGTTAATTAATGATCCCAT-----GCGTAAATTTTACGCAAC 1009
| | | | |
Db 961 TAAAGTGTATCAGTGTGAGATCACGCTTTCAGTGAAGCAAGAAATTTGATTTGATTA 1020
| | | | |
QY 1010 TATCTTTTGAAGGTTAATCATGCTGCATGAGATCATATCGTGGGCTTTTTTCCGGCT 1069
| | | | |
Db 1021 TTTTCGAGAGAAAGATGCGAGAAATCAATACATTAAGATTTGAATCTTTTGTGTCT 1080
| | | | |
QY 1070 CAGTCAATCCCGCAAGTGTGCGCTATCTGGGCAATCGGAGAGAAAGAACCCGCTGCTTTTC 1129
| | | | |
Db 1081 TCGACATGGGTATCTCATGTGTGAATGGCCCTAAGAGATCCGCGCAAGCTTGCATGCG 1140
| | | | |
QY 1130 CCGGAGGTTTGAAGCGGATGGAAGATTTGCCAGAGATGAAGTGTGTGATTTGAGGT 1189
| | | | |
Db 1141 TGAAGTGTGACTCTAAGAGATCCCGAGCGCTGAGCGCATTAATATATGTTTTCCGTAAT 1200
| | | | |
QY 1190 TGAAGCAAAACGACGTTTACATGATGATTCGGGAAGGTGTGGCATGACGCTTTAA 1249
| | | | |
Db 1201 TAGGCGCTTCCATGATGAGACAGGCGCTT--GATGTGACGGGATGAACATTAATAG 1258
| | | | |

QY 1250 CGGTAACTGTTCGTTCAAGGCGACCTGGAGTACCAGTTGTGCGGCTTTCCGACACA 1309
| | | | |
Db 1259 CAATGACGAGCAATTAATCTCAACAG--AGCAGAAAACGAGGCTATCCACAAGT 1315
| | | | |
QY 1310 GTTCCGAGTGTGAGCCCGAAGCGCATCAAGCAACCCGAAACAATACCGGCGACG----- 1363
| | | | |
Db 1316 CAGCGTACCATTAAGCGCAAGCCTCAACGACGAGCAGACACGAGCGGTGAGTGA 1375
| | | | |
QY 1364 CCGAACTGCGCGCGGTGTGAGATTATGACAGCGGTGCGGCGCTGGATTTAGCT 1423
| | | | |
Db 1376 TCCAAATCTTTTACTGTGAGAAATCGAAATCATCTTGTGTTAATCAAAACGCGAG 1435
| | | | |
QY 1424 CAGCGAGACGGGTAT-----CTGTGATGATCCGAGAAATGACATGATATCCCGAG 1479
| | | | |
Db 1436 AAGCTGAATGAGAAATTCGACCTCGAGGGGGGCGCGATCCAGCTTTGTTCCCTTTA 1495
| | | | |
QY 1480 AGTTACCGCGCGCGCGCTTGGCGTATCATGTGTATGCTGTGTTCTGTGTGAATTG 1539
| | | | |
Db 1496 GTGAGGTTAATTCGAGCTTGGCGTATCATGTGTATGCTGTGTTCTGTGAAATTG 1555
| | | | |
QY 1540 TTATCCGCTCACAATTCACACACAATACAGACCGGAAAGCATTAAGTGAAGCCTGGGG 1599
| | | | |
Db 1556 TTATCCGCTCACAATTCACACACAATAGAGCCGGAAGCATTAAGTGAAGCCTGGGG 1615
| | | | |
QY 1600 TGCCTATGATGAGCTAACTCAATTAATGCGTTGCGCTCACTGCGCGCTTTCAGTTC 1659
| | | | |
Db 1616 TGCCTATGATGAGGATTAATCAATTAATGCGTTGCGCTCACTGCGCGCTTTCAGTTC 1675
| | | | |
QY 1660 GGGAAACCTGTGTCGACAGCTGCAATTAATGATTCGCGCAACGCGCGGAGAGAGCGGTTT 1719
| | | | |
Db 1676 GGGAAACCTGTGTCGACAGCTGCAATTAATGATTCGCGCAACGCGCGGAGAGAGCGGTTT 1735
| | | | |
QY 1720 GCGTATTTGGCGCTTTCGCTTCCGCTCACTGACTCGCTGCGCTGCGTTCGCT 1779
| | | | |
Db 1736 GCGTATTTGGCGCTTTCGCTTCCGCTCACTGACTCGCTGCGCTGCGTTCGCT 1795
| | | | |
QY 1780 GGGGAGCGGATTAAGCTCACTCAAGCGGTAATCGTTATTCACAGATCAGAGGA 1839
| | | | |
Db 1796 GGGGAGCGGATTAAGCTCACTCAAGCGGTAATCGTTATTCACAGATCAGAGGA 1855
| | | | |
QY 1840 TAAAGCAAGAAAGAAATGTGAGCAAGAAAGGCGCAAGAAAGGCGGAAACCGTAAAGAGC 1899
| | | | |
Db 1856 TAAAGCAAGAAAGAAATGTGAGCAAGAAAGGCGCAAGAAAGGCGGAAACCGTAAAGAGC 1915
| | | | |
QY 1900 GCGGTTGCTGCGGTTTTTCCATAGAGCTCCGCGCTGACAGAGATCAAAATTCAGC 1959
| | | | |
Db 1916 GCGGTTGCTGCGGTTTTTCCATAGAGCTCCGCGCTGACAGAGATCAAAATTCAGC 1975
| | | | |
QY 1960 CTCAAGTGAAGGAGGAGGAGAAACCGGACAGAGCTTAAGATTAAGAGCTTCCCGCTGG 2019
| | | | |
Db 1976 CTCAAGTGAAGGAGGAGGAGAAACCGGACAGAGCTTAAGATTAAGAGCTTCCCGCTGG 2035
| | | | |
QY 2020 AAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGACCGCTTAACGGATACCGTCCGCTT 2079
| | | | |
Db 2036 AAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGACCGCTTAACGGATACCGTCCGCTT 2095
| | | | |
QY 2080 TCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGATGATCTCAAGTTGAGT 2139
| | | | |
Db 2096 TCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGATGATCTCAAGTTGAGT 2155
| | | | |
QY 2140 GTAAGTGTGCTGCTCAAGAGCGGAGCTGTGACAGAAACCCCGCTTACGCGGACGCGCTG 2199
| | | | |
Db 2156 GTAAGTGTGCTGCTCAAGAGCGGAGCTGTGACAGAAACCCCGCTTACGCGGACGCGCTG 2215
| | | | |
QY 2200 CGCTTAATCCGATTAATCATGCTTTGATGCTCAACCCGCTTAAGACAGCATTAATCGCACT 2259
| | | | |
Db 2216 CGCTTAATCCGATTAATCATGCTTTGATGCTCAACCCGCTTAAGACAGCATTAATCGCACT 2275
| | | | |
QY 2260 GGCAGAGCAACTGTGTAACAGATTTAGCAGAGGAGATGAGCGGTGCTACAGAGTT 2319
| | | | |
Db 2276 GGCAGAGCAACTGTGTAACAGATTTAGCAGAGGAGATGAGCGGTGCTACAGAGTT 2335
| | | | |
QY 2320 CTGAAGTGTGCGCTTAATCAAGGCTACATGAAGAGAGATTTTGTGATCTGCGCTCT 2379
| | | | |

Db 2336 CTTGAAGGTGGCTTAACCTACGCGCTACAGTAAGAGCAAGTATTTGGTATCTGGCGCTCT 2395
Qy 2380 GCTGAAGCACTTAACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGGCAACCAACAC 2439
Db 2396 GCTGAAGCACTTAACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGGCAACCAACAC 2455
Qy 2440 CGCTGTAAGCGGTGTTTTTTTTTTTGGTGAACAGCAAGTTACGGCGAAGAAAAAGATC 2499
Db 2456 CGCTGTAAGCGGTGTTTTTTTTTTTGGTGAACAGCAAGTTACGGCGAAGAAAAAGATC 2515
Qy 2500 TCAAGAGATCTTTGATCTTTCTAAGGGGCTGACGCTGAGTGAAGCAAGAACTGACG 2559
Db 2516 TCAAGAGATCTTTGATCTTTCTAAGGGGCTGACGCTGAGTGAAGCAAGAACTGACG 2575
Qy 2560 TTAAGGATTTTGGTCATGAGATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTA 2619
Db 2576 TTAAGGATTTTGGTCATGAGATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTA 2635
Qy 2620 AAAATGAAGTTTAAATCTAATCTAAGATATATGAGTAACTTGGTCTGACAGTTACCA 2679
Db 2636 AAAATGAAGTTTAAATCTAATCTAAGATATATGAGTAACTTGGTCTGACAGTTACCA 2695
Qy 2680 ATGCTTAATCAGTAGAGGCACTTATCTGAGCGATCTGTATTTGGTTCATCATATAGTTGC 2739
Db 2696 ATGCTTAATCAGTAGAGGCACTTATCTGAGCGATCTGTATTTGGTTCATCATATAGTTGC 2755
Qy 2740 CTGAATCCCGCTGCTGATAGATTAACGATAAGGAGGCTTACATCTGAGCCCAAGTGC 2799
Db 2756 CTGAATCCCGCTGCTGATAGATTAACGATAAGGAGGCTTACATCTGAGCCCAAGTGC 2815
Qy 2800 TGCATATGATACCGGAGAACCAAGCTCACCGGCTTCAGATTTATCAGCAATTAACCAAGC 2859
Db 2816 TGCATATGATACCGGAGAACCAAGCTCACCGGCTTCAGATTTATCAGCAATTAACCAAGC 2875
Qy 2860 AGCGGAAAGGCGCAGAGCGCAGAGTGTCTGCAACTTATCCGCTCATCATCAGTCTAT 2919
Db 2876 AGCGGAAAGGCGCAGAGCGCAGAGTGTCTGCAACTTATCCGCTCATCATCAGTCTAT 2935
Qy 2920 TAAATTTGTCGGGAAAGCTAGAGTAAAGTTCGCAAGTAAATAGTTTGCACAAGTGT 2979
Db 2936 TAAATTTGTCGGGAAAGCTAGAGTAAAGTTCGCAAGTAAATAGTTTGCACAAGTGT 2995
Qy 2980 TGCATATGATACCGGAGAACCAAGCTCACCGGCTTCAGATTTATCAGCAATTAACCAAGC 3039
Db 2996 TGCATATGATACCGGAGAACCAAGCTCACCGGCTTCAGATTTATCAGCAATTAACCAAGC 3055
Qy 3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCGCCATGTTTGCAAAAAAGGGTTAG 3099
Db 3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCGCCATGTTTGCAAAAAAGGGTTAG 3115
Qy 3100 CTCCTTCGGTCTCCGATCGTTGTACAGAGTAAAGTTGGCCGACAGTGTATCATCATAGT 3159
Db 3116 CTCCTTCGGTCTCCGATCGTTGTACAGAGTAAAGTTGGCCGACAGTGTATCATCATAGT 3175
Qy 3160 TATGCGACGATGATTAATCTCTTACTGTATGATCCGTAAAGATGCTTTTCTGTGAC 3219
Db 3176 TATGCGACGATGATTAATCTCTTACTGTATGATCCGTAAAGATGCTTTTCTGTGAC 3235
Qy 3220 TGTGAGATCTCAACCAAGTATCTGAGATAGTATGAGTGGGAGACCGAGTCTCTGTG 3279
Db 3236 TGTGAGATCTCAACCAAGTATCTGAGATAGTATGAGTGGGAGACCGAGTCTCTGTG 3295
Qy 3280 CCGGCGCTCAATACGGGATTAATACCGGCGCATATAGCAAGAACTTAAAGTGTCTATCAT 3339
Db 3296 CCGGCGCTCAATACGGGATTAATACCGGCGCATATAGCAAGAACTTAAAGTGTCTATCAT 3355
Qy 3340 TGGAAACGTTCTTCGGGCGCAAACTTCAAGATCTTACCGCTGTGAGATCAAGTTC 3399
Db 3356 TGGAAACGTTCTTCGGGCGCAAACTTCAAGATCTTACCGCTGTGAGATCAAGTTC 3415
Qy 3400 GATGTAACCACTGCTGACCACTGATCTTACAGCACTTTTCACTTACAGAGTTTC 3459

Db 3416 GATGTAACCACTGCTGACCACTGATCTTACAGATCTTTTACACAGGTTTC 3475
Qy 3460 TGGGTGAGCAAAAAACAGAAAGCAAAATGCGGCAAAAAAGGAAATTAAGGCGACACGAA 3519
Db 3476 TGGGTGAGCAAAAAACAGAAAGCAAAATGCGGCAAAAAAGGAAATTAAGGCGACACGAA 3535
Qy 3520 ATGTTGAATCTCATCT 3579
Db 3536 ATGTTGAATCTCATCT 3595
Qy 3580 TCTCATGAGCGGATACATATTTGAATGATTTGAATAAATAAACAATAAGGGGTTCCGCG 3639
Db 3596 TCTCATGAGCGGATACATATTTGAATGATTTGAATGATTTGAATAAATAAACAATAAGGGGTTCCGCG 3655
Qy 3640 CACATTTCCCGGAAAGTCCAC 3662
Db 3656 CACATTTCCCGGAAAGTCCAC 3678

RESULT 7
A20700 LOCUS A20700 5314 bp DNA linear PAT 22-AUG-1994
DEFINITION pMTV1 DNA sequence.
ACCESSION A20700
VERSION A20700.1 GI:583402
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1 (bases 1 to 5314)
AUTHORS
TITLE
CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
USE AS AN IMMUNOGEN AND VACCINE
JOURNAL
Patent: WO 913155-A 4 05-SEP-1991;
FEATURES
source
1..5314
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 71.6%; Score 2623.6; DB 6; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;
Qy 20 ATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGTCAATTTTAAACCAATAGGCC 79
Db 1651 ATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGTCAATTTTAAACCAATAGGCC 1710
Qy 80 GAAATCGCAAAATCCCTTATTAATCAAAAGATGACCGAGATAGGGTTGAGTGTGT 139
Db 1711 GAAATCGCAAAATCCCTTATTAATCAAAAGATGACCGAGATAGGGTTGAGTGTGT 1770
Qy 140 CCAATTTGAAACAAAGATCCACTTAATTAAGAACTGGAATCCCAAGTCAAAAGGCGAAAA 199
Db 1771 CCAATTTGAAACAAAGATCCACTTAATTAAGAACTGGAATCCCAAGTCAAAAGGCGAAAA 1830
Qy 200 ACCGTATACGAGGCGATGAGCCCACTACGTGAACCATCATCCCTATACAGTTTGTGGGG 259
Db 1831 ACCGTATACGAGGCGATGAGCCCACTACGTGAACCATCATCCCTATACAGTTTGTGGGG 1890
Qy 260 TCGAGTGGCCGTAAGACATTAATCGGAACCTTAAGAGGAGCCCGCATTTAGACTTGA 319
Db 1891 TCGAGTGGCCGTAAGACATTAATCGGAACCTTAAGAGGAGCCCGCATTTAGACTTGA 1950
Qy 320 CCGGGAAGGCGGGAACGTGGCGGAAGAAAGGGAAGAAAGGGAAGGCGGGCGCT 379
Db 1951 CCGGGAAGGCGGGAACGTGGCGGAAGAAAGGGAAGAAAGGGAAGGCGGGCGCT 2010
Qy 380 AGGCGCTGCAAGTATGAGGCTACAGCTCGCGTAAACCAACCAACCGCGCGCTTAAT 439
Db 2011 AGGCGCTGCAAGTATGAGGCTACAGCTCGCGTAAACCAACCAACCGCGCGCTTAAT 2070

QY 440 GCGCGCTACAGAGGCGCGCATTCGCCATTGAGGCTGCCAATCTGTTGGAGAGGCGCA 499
|
|
|
Db 2071 GGGCGGCTACAGAGGCGCGCTCCCATTCGCCATTGAGGCTGCCAATCTGTTGGAGAGGCGCA 2130
|
|
|
QY 500 TCGGTGCGGCGCTCTTCGCTATTACGCGAGCTGGGAGGGGAGATGTGTGCAAGGCGCA 559
|
|
|
Db 2131 TCGGTGCGGCGCTCTTCGCTATTACGCGAGCTGGGAGGGGAGATGTGTGCAAGGCGCA 2190
|
|
|
QY 560 TTAAGTTGGGTAAAGCCGACAGGCTTTCCAGTCAAGAGCTGTGTAACGACGAGCGCATGAG 619
|
|
|
Db 2191 TTAAGTTGGGTAAAGCCGACAGGCTTTCCAGTCAAGAGCTGTGTAACGACGAGCGCATGAA 2250
|
|
|
QY 620 CGCGCGCTGCTTCAATTCAGGTTTTGAAACCCGAGGAGAGCGGAGAGAGCTGCGGGGTGAAT 679
|
|
|
Db 2251 TTGTAAATACGACTCATATAGGGGCAATTGGAAGCTCCACCGCGGTGGCGGCGCTTAAT 2310
|
|
|
QY 680 GTGTTTACAGCGGTATGAGAGCAATGTAAGATGTCTGACAGCTGACGAACACGACAGCTA 739
|
|
|
Db 2311 ATGTGCACTCTCAATCAATCTGCTGATGCGGACATGTTAAGCCAGTATATACATC 2370
|
|
|
QY 740 GATTAAACCTAGAAAGATATCATATGTTGACGTACGTAAAGATATCATGCGTAAAT 799
|
|
|
Db 2371 CGGTATGCTACGTGACTGGGTCAATGGCTGCGCCGACACCGCCAAACCGCGCTGACG 2430
|
|
|
QY 800 TGACGCATGGGATCTGTAAATACGACTCATATAGGGGCAATTGGGTACCGGCGCCCGCT 859
|
|
|
Db 2431 CGCCCTGACGGGCTGTCTGTCTCCGGGATCCGCTTACAGACAGCTGTACCGCTCTCCG 2490
|
|
|
QY 860 CGAGGTGAGCGGTATGCA-----TAAAGCTGATATGAAATTCCTGCAAGCCGCGGGATC 913
|
|
|
Db 2491 GGAAGTCAATGTGTCAAGAGGTTTTACCGTATCAACGGAACGCGGACAGAGCAATGAGTC 2550
|
|
|
QY 914 CACTAGTTCTAGAGCGGCGCCGACGCGGTGAGGCTCCAGCTTTTGTTCCTTTAATGAGAG 973
|
|
|
Db 2551 GGAATGCTTTGAGCAATTCGTCTTAACTTAAGTAAGCAATTCGTAAAGCTGTCACTGTGC 2610
|
|
|
QY 974 GGTAAATTAAGTCCCAT-----GCGTCAATTTTACGACACTATCTTTCTAGGGTTAA 1026
|
|
|
Db 2611 GGAATCACCGCTTCAGTAGGAGCAAGAAAGCAATGATGTAAATTTGAGAGAAAGATG 2670
|
|
|
QY 1027 TCTAGCTGCATCAGGATCATATGCTGGGGTCTTTTCCGGCTCATGCTGCGCCAAAGCT 1086
|
|
|
Db 2671 CGAGGAAGATCATATACATAAGAGTTGAATCTTGTGTGTTGACATGAGGTAAATCTCT 2730
|
|
|
QY 1087 GGGGCTATCTGGGCATCGGGAGAGAAAGACCCGTCCTTTTCCCGCAGGTTGAAGCG 1146
|
|
|
Db 2731 CATGTTGAATGGCCCTAGAGATCGGCGCAAGCTTGCAATGCTGAGGTGCACTTA 2790
|
|
|
QY 1147 CATGAAAGAGTTTGGCGAGATGATGCTGCTGCAATTGACGTTGAGCGAAAACGACGT 1206
|
|
|
Db 2791 GGAATCCCGACGCTCGACGCCATTAATTAATGTTTCCGTAAATTCAGCGCTTCCATGAT 2850
|
|
|
QY 1207 TTAACATGATGATTCGGGAAGTGTGSCATGCAAGCTTTTAAACGAGTGAATGTTGCTTC 1266
|
|
|
Db 2851 GAGACAGGCGGTTT--GAATGTTGAGCGGATGAAACATAATAGCAATGACGCGACATA 2908
|
|
|
QY 1267 AGGCGACCTGGGATACCAATGCTGCGGCTTTTCCGACACAGTTCCGAGATGTACAGCC 1326
|
|
|
Db 2909 AACTCAACAGG---ACGAGAAAGCGAGGGTATCCCAAAAGTCCAGCGTACATTAACG 2965
|
|
|
QY 1327 CGAAGGCATCAGCAACCGGAACAATACGSGGACAG-----CGGAATCCGCTGCG 1380
|
|
|
Db 2966 CAAAGCTCAACGAGAGAGAGCAAGAGCGGTACAGTCAATCCAACTTTGTTACTTC 3025
|
|
|
QY 1381 GTGTGCAATTAATGACAGCGGTGCGCGCTGGGATATTAAGTCAAGAGAGAGCGGTAT- 1439
|
|
|
Db 3026 GTTCAGAAATTCAGAAATCATCTTCGTTAAATCCAAACGCGAGAGCTGTAATGAAAT 3085
|
|
|
QY 1440 ---CCGTGCTGATGCGGAGAAATGGAATAGATATCCCGTGAATACCGGCGGGCGC 1496
|
|
|
Db 3086 CGAAGCTCAAGGGGGGCGGTAACCAAGCTTTGTTCCCTTAAATGAGGGTTAATTCGA 3145
|
|
|
QY 1497 GCTTGGCGTAAATCATGCTATAGCTGTTTCTGTGAAATTTGTAATCCGCTCACAATTC 1556
|
|
|

Db 3146 GCTTGGGCTAATATGCTCATACCTGTTTCTGTGAAATTTTATCCGCTACAAATTC 3205
|
|
|
QY 1557 CACACAAATATACAGCGCGGAAGCATTAAGTAAAGCTTGGGGTGCCTAATGATGAGCT 1616
|
|
|
Db 3206 CACACAAATATGAGAGCGGGAAGCATTAAGTAAAGCTTGGGGTGCCTAATGATGAGGT 3265
|
|
|
QY 1617 AACTCAATTAATTTGGGTCGCTCACTGCGGCTTTCCAGTCGGAACCTGTCTGCC 1676
|
|
|
Db 3266 AACTCAATTAATTTGGGTCGCTCACTGCGGCTTTCCAGTCGGAACCTGTCTGCC 3325
|
|
|
QY 1677 AGCTCAATTAATGATATGCGCAACCGCGCGGAGAGAGCGGTTTGCATATGGAGCGCTCT 1736
|
|
|
Db 3326 AGCTCAATTAATGATATGCGCAACCGCGCGGAGAGAGCGGTTTGGGATATGGGCGCTCT 3385
|
|
|
QY 1737 CGGCTTCTCGCTCACTGACTGCTGCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1796
|
|
|
Db 3386 CGGCTTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445
|
|
|
QY 1797 CTCACCTAAAGCGGTAAATACGTTATCCACAGAAATCAGAGGATTAACGAGAAAGACA 1856
|
|
|
Db 3446 CTCACCTAAAGCGGTAAATACGTTATCCACAGAAATCAGAGGATTAACGAGAAAGACA 3505
|
|
|
QY 1857 TGTGAGCAAAAGCCGACAAAGGCGCAGAAACCTTAATAAGCGCGGCTTGTGCGGTTTT 1916
|
|
|
Db 3506 TGTGAGCAAAAGCCGACAAAGGCGCAGAAACCTTAATAAGCGCGGCTTGTGCGGTTTT 3565
|
|
|
QY 1917 TCCATAGGCTCGGCGCCCTGACGAGCATCAAAATTCAGCGCTCAAGTACAGAGTGGC 1976
|
|
|
Db 3566 TCCATAGGCTCGGCGCCCTGACGAGCATCAAAATTCAGCGCTCAAGTACAGAGTGGC 3625
|
|
|
QY 1977 GAAACCGACAGGACTAATTAAGATACAGAGGTTTTCCCTGGAAGCTCCCTGCGCT 2036
|
|
|
Db 3626 GAAACCGACAGGACTAATTAAGATACAGAGGTTTTCCCTGGAAGCTCCCTGCGCT 3685
|
|
|
QY 2037 CTCGTTCCGACCGTCGCGGCTTACCGGATACGTCGCTGCTTCTCCCTTCCGGAAGCG 2096
|
|
|
Db 3686 CTCGTTCCGACCGTCGCGGCTTACCGGATACGTCGCTGCTTCTCCCTTCCGGAAGCG 3745
|
|
|
QY 2097 TGGCGCTTCTCATATGCTCAGCTGTAGATCTCAGTTCGGTGAAGGTGCTGCTCA 2156
|
|
|
Db 3746 TGGCGCTTCTCATATGCTCAGCTGTAGATCTCAGTTCGGTGAAGGTGCTGCTCA 3805
|
|
|
QY 2157 AGCTGGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGCGCTTATCGGTAAT 2216
|
|
|
Db 3806 AGCTGGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGCGCTTATCGGTAAT 3865
|
|
|
QY 2217 ATGCTTTGAATCCAAACCGGTAAAGACAGACTTATGCGCATGSGCAGGCGCACTGTA 2276
|
|
|
Db 3866 ATGCTTTGAATCCAAACCGGTAAAGACAGACTTATGCGCATGSGCAGGCGCACTGTA 3925
|
|
|
QY 2277 ACAGATTTAGAGAGCGAGATATGATAGGCGGTGTCAGAGTTCTTGAATGTGTGCTTA 2336
|
|
|
Db 3926 ACAGATTTAGAGAGCGAGATATGATAGGCGGTGTCAGAGTTCTTGAATGTGTGCTTA 3985
|
|
|
QY 2337 ACTAGGCTACACTAAGAAAGACGTAATTTGTAATTCGCGCTTCTGCTGAAGCAGTTACT 2396
|
|
|
Db 3986 ACTAGGCTACACTAAGAAAGACGTAATTTGTAATTCGCGCTTCTGCTGAAGCAGTTACT 4045
|
|
|
QY 2397 TCGGAAAAAGAGTTGATGCTTGTGATCCGGCAACAAACCAACCGCTGATAGGCGGT 2456
|
|
|
Db 4046 TCGGAAAAAGAGTTGATGCTTGTGATCCGGCAACAAACCAACCGCTGATAGGCGGT 4105
|
|
|
QY 2457 TTTTGTGTTTCAAGCAGCAATTAACGCGCAAAAGAAAGAGATCAAGAAAGATCCTTTGA 2516
|
|
|
Db 4106 TTTTGTGTTTCAAGCAGCAATTAACGCGCAAAAGAAAGAGATCAAGAAAGATCCTTTGA 4165
|
|
|
QY 2517 TCTTTTCTACGGGATCTGACGCTCAGTGAAGAGAAACCAACGTTAAAGGATTTTGTCA 2576
|
|
|
Db 4166 TCTTTTCTACGGGATCTGACGCTCAGTGAAGAGAAACCAACGTTAAAGGATTTTGTCA 4225
|
|
|
QY 2577 TGAATTAACAAAGAGATCTTCACTTAATCTTTTAAATTAATAATGAAGTTTAAAT 2636
|
|
|

Db 4226 TGAATATTCAAAAAGATCTTCACTAGATCCITTTAAATTAATAAGATTTTAAT 4285
Qy 2637 CAATCTAAAGTATATATGAGTAAACTTGCTGACAGTTACCAATGCTTAATCACTGAGG 2696
Db 4286 CAACTAAAGATATATATGAGTAAACTTGCTGACAGTTACCAATGCTTAATCACTGAGG 4345
Qy 2697 CACCATCTGAGCATCTGCTATTTTGCTCATCATGTTGGCCGATCTCCCGCCGCTGT 2756
Db 4346 CACCATCTGAGCATCTGCTATTTTGCTCATCATGTTGGCCGATCTCCCGCCGCTGT 4405
Qy 2757 AGATACATACGATACGGAGAGGCTTACCATCTGAGGCTGAGTGCATATGATACCGAGAG 2816
Db 4406 AGATACATACGATACGGAGAGGCTTACCATCTGAGGCTGAGTGCATATGATACCGAGAG 4465
Qy 2817 ACCCAAGCTCAACCGGCTCCAGATTTATCAGCAATTAACCAAGCCAGCCGAAAGGCGAGC 2876
Db 4466 ACCCAAGCTCAACCGGCTCCAGATTTATCAGCAATTAACCAAGCCAGCCGAAAGGCGAGC 4525
Qy 2877 GCAGAAAGTGTCCGCAACTTTATCCGCTCCATCCCACTCTTAAATTTGTTGCCGGAGAG 2936
Db 4526 GCAGAAAGTGTCCGCAACTTTATCCGCTCCATCCCACTCTTAAATTTGTTGCCGGAGAG 4585
Qy 2937 CTAGAGTACATAGTTCGCACTTAATAGTTTGCAGCAAGTTGTCATTTGCTACAGGCA 2996
Db 4586 CTAGAGTACATAGTTCGCACTTAATAGTTTGCAGCAAGTTGTCATTTGCTACAGGCA 4645
Qy 2997 TCGTGTGTCAAGCTCGTGTGTTGGTATGAGCTTCAATCAGCTCCGGTTCCCAAGATCAA 3056
Db 4646 TCGTGTGTCAAGCTCGTGTGTTGGTATGAGCTTCAATCAGCTCCGGTTCCCAAGATCAA 4705
Qy 3057 GGCAGATTAATGATATCCCCCAATGTTGTCAGAAAAAGCGGTTAGTCTCTCCGCTCTCCGA 3116
Db 4706 GGCAGATTAATGATATCCCCCAATGTTGTCAGAAAAAGCGGTTAGTCTCTCCGCTCTCCGA 4765
Qy 3117 TCGTGTTCAGAAATAGTTGAGCCGAGTGTATCATCTAGTATGAGCAGACATGCTATA 3176
Db 4766 TCGTGTTCAGAAATAGTTGAGCCGAGTGTATCATCTAGTATGAGCAGACATGCTATA 4825
Qy 3177 ATTCTCTTACTGTCAATGCTATCCGTAAGATGCTTTTCTGTGATCTGTGATGATCTCAACCA 3236
Db 4826 ATTCTCTTACTGTCAATGCTATCCGTAAGATGCTTTTCTGTGATCTGTGATGATCTCAACCA 4885
Qy 3237 AGTCAATCTGAGAAATAGTGTATGCGGCGAGCCGAGTGTCTTTCGCGCGGCTCAATACCGG 3296
Db 4886 AGTCAATCTGAGAAATAGTGTATGCGGCGAGCCGAGTGTCTTTCGCGCGGCTCAATACCGG 4945
Qy 3297 ATAAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGATGAGAAAAAGTTCTTGG 3356
Db 4946 ATAAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGATGAGAAAAAGTTCTTGG 5005
Qy 3357 GGGCAAAAATCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGATCACTCGTG 3416
Db 5006 GGGCAAAAATCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTGATGATCACTCGTG 5065
Qy 3417 CACCAATCTGATCTTACGATCTTTTACTTCAACAGGTTTCTGGGAGACAAAAACAG 3476
Db 5066 CACCAATCTGATCTTACGATCTTTTACTTCAACAGGTTTCTGGGAGACAAAAACAG 5125
Qy 3477 GAAAGCAAAAATGCGCAAAAAGGAAATTAAGGCGACACAGGAAATGTTGAATATCTCATAC 3536
Db 5126 GAAAGCAAAAATGCGCAAAAAGGAAATTAAGGCGACACAGGAAATGTTGAATATCTCATAC 5185
Qy 3537 TCTTCTCTTTTCAATATTTATGAAAGATTTATCAGGTTATTTGTCTCATGAGCGATACA 3596
Db 5186 TCTTCTCTTTTCAATATTTATGAAAGATTTATCAGGTTATTTGTCTCATGAGCGATACA 5245
Qy 3597 TATTTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTTCCCGGAAAG 3656
Db 5246 TATTTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTTCCCGGAAAG 5305
Qy 3657 TGCCAC 3662
Db 5306 TGCCAC 5311

RESULT 8
LOCUS 115649 5314 bp DNA linear PAT 02-APR-1996
DEFINITION Sequence 4 from parent US 5470573.
ACCESSION 115649
VERSION 115649.1 GI:1250557
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5314)
AUTHORS Lubitz,W. and Sostak,M.P.
TITLE Immunogens comprising the non-lytic membrane spanning domain of bacteriophages MS2 or PhiX174
JOURNAL Patent: US 5470573-A 4 28-NOV-1995;
FEATURES
LOCATION/Qualifiers
1..5314
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 71.6%; Score 2623.6; DB 6; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;
Qy 20 ATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTAAACCAATAGGCC 79
Db 1651 ATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTAAACCAATAGGCC 1710
Qy 80 GAAATCGGCAAAATCCCTTAAATCAAAGAATAGACCGAGATAGGAGTTGAGTTGTT 139
Db 1711 GAAATCGGCAAAATCCCTTAAATCAAAGAATAGACCGAGATAGGAGTTGAGTTGTT 1770
Qy 140 CCAATTTTGAACAAGATGCCATTTAAAGACGTGACTCCAGCTCAAGGCGAAAA 199
Db 1771 CCAATTTTGAACAAGATGCCATTTAAAGACGTGACTCCAGCTCAAGGCGAAAA 1830
Qy 1771 CCAATTTTGAACAAGATGCCATTTAAAGACGTGACTCCAGCTCAAGGCGAAAA 1830
Qy 200 ACCGTTATACAGGCGATAGGCCCACTAGTAAACCATCACTCAATCAAGTTTGGGG 259
Db 1831 ACCGTTATACAGGCGATAGGCCCACTAGTAAACCATCACTCAATCAAGTTTGGGG 1890
Qy 260 TCGAGGTGCGGTAAAGCACTTAATCGGAACCTTAAAGGAGCCCCGATTTAGCTTGA 319
Db 1891 TCGAGGTGCGGTAAAGCACTTAATCGGAACCTTAAAGGAGCCCCGATTTAGCTTGA 1950
Qy 320 CCGGGAAGGCGCGGAAAGTGGCGGAGAAAGGAGGAAAGGAGGAGGCGGCGCT 379
Db 1951 CCGGGAAGGCGCGGAAAGTGGCGGAGAAAGGAGGAAAGGAGGAGGCGGCGCT 2010
Qy 380 AGGCGCTGAGCAAGTGTAGGCGTCAAGCTGCGCTAACAACACACCGCGCGCTTAAT 439
Db 2011 AGGCGCTGAGCAAGTGTAGGCGTCAAGCTGCGCTAACAACACACCGCGCGCTTAAT 2070
Qy 440 GCGCGCTTACAGGCGCGCTCCCATTTGCCATTTACAGCTTGGCGCAACTGTTGGAAAGGCGGA 499
Db 2071 GCGCGCTTACAGGCGCGCTCCCATTTGCCATTTACAGCTTGGCGCAACTGTTGGAAAGGCGGA 2130
Qy 500 TCGGTGGGGGCTTTGCTTATTAAGCGCAGCTGCGGAAAGGGGAGTGTGTCGCAAGGCGGA 559
Db 2131 TCGGTGGGGGCTTTGCTTATTAAGCGCAGCTGCGGAAAGGGGAGTGTGTCGCAAGGCGGA 2190
Qy 560 TTAAGTGGGTAAAGCGCAGGAGTTTCCAGTCAAGAGCTTGTAAACGAGCGGCGAGTGA 619
Db 2191 TTAAGTGGGTAAAGCGCAGGAGTTTCCAGTCAAGAGCTTGTAAACGAGCGGCGAGTGA 2250
Qy 620 CGCGCTCGTTCAATTCAGTTTGAACCGGTGAGAGCGGCGAGACTCGCGGTGCAAT 679
Db 2251 TTGTAAATACGATCTCATATAGGCGGAATGGAAGTTCACCGCGGTGCGCGCTTAGT 2310
Qy 680 GTGTTTACAGCGATGAGCAGATGAAGTGTGACACGCTGAGAAACAGCAGACTA 739

Db 2311 ATGTGCACTCTCAGTACATCTGCTGTGATGCCGATAGTTAAGCCAGTATATACATC 2370
Qy 740 GATTAAACCTAGAAAGTATCATATTTGTCAGTCATTTAAAGATATCATGCTAAAT 799
Db 2371 CGCTATGCTACGTACATGAGTATGCTGCGCCGACACCCGCAACACCCGCTGACG 2430
Qy 800 TGAAGCATGAGATCTGTAAATACGATCATATAGGCGAAATTTGGTACCGGGCCCCCT 859
Db 2431 CGCCCTGACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAAGCTGTACCTCTCG 2490
Qy 860 CAGAGCTCAGCGATATGA-----TAACTGTATATCGAAATTCCTGACGCCCGGGAGATC 913
Db 2491 GGAAGTCATGTGTCAAGAGTTTTCACCCCTCATCACGAAACGCCGAGGCGATAAAGTTC 2550
Qy 914 CACTAGTCTTAGAGCGCGCCGACCGCGGTGAGCTCCAGCTTTTGTTCCTTTAGGTAA 973
Db 2551 GGAATGCTTTGTGACAAATTCCTCTTAAAGTAAATGCTGTAAATGCTGCTACTGTC 2610
Qy 974 GGTAAATTAGATCCAT-----GCTCAATTTTACGACATATCTTTCTAAGGTTAA 1026
Db 2611 GGATCACCGCTTCCAGTACGACAGAAAGCAATTTGTTAAATTTCCAGAAAGATCG 2670
Qy 1027 TCTAGCTGATCAGATATATCTCGGATCTTTTCCGCTCATGTCATGCCCAACT 1086
Db 2671 CGAGAAATCATATCATTAAGATTGAATCTTTTGTGTCTTGCATGAGTATATCT 2730
Qy 1087 GAGCTATCTGGCATCGGAGGAGAAAGCCCGCTTTTCCCGAGGTTGAACCG 1146
Db 2731 CATGTTTAAATGGCCCTTAGAGGATCCGGCCAAAGCTTGATGCTGACAGTCTAAGA 2790
Qy 1147 CATGAAAGATTGGCCAGATGATCTCTGATTTAGCTTGAACGAAACGCACT 1206
Db 2791 GGATCCCGACGCTGACGCAATTAATATGTTTCCGTAAATTCACGCTTCATGAT 2850
Qy 1207 TTACCATGATGATTCGGAGATGTGCGCATGACAGCTTTTAAAGGTAATCTGTC 1266
Db 2851 GAGACAGGCGCTTT--GAATGTACCGGATGAACATTAATGAATGACGCGACATA 2908
Qy 1267 AGGCACCTGGGATACCAATTCGTGCGGCTTTTCCGGAACAAGTTCGAGTGTACGC 1326
Db 2909 AACTCAACAG--AGCAGAAAGCGAGGATATCCCAAAATCCAGCTGATCAATAAGC 2965
Qy 1327 CGAAGCATCAGCAACCCGAACAATACCGGCAAG-----CCGGAATTCGCGCGC 1380
Db 2966 CAACCTCTACGACGCAACGACGAGACGCTGATGCAATCCAAATTTGTATATCTC 3025
Qy 1381 GTGTGCAATTAATGACAGCGGTGCGCGCTGGGATTTATCTCAGCGAGACGAGTAT- 1439
Db 3026 GTCAAAATCGAATCATCTCGGTTAAATCCAAACGCGCAAAACCTGAAATGAATTT 3085
Qy 1440 ---CCTGGCTGAGTCCGCGAAGAAATGACATGATACCCCGTGAATTTCCCGCGGCGC 1496
Db 3086 CGACTCTGAGGGGGGCCCCGGTACCCAGCTTTTGTTCCTTTAGTAGGTTAAATCCGA 3145
Qy 1497 GCTTGGCTAATCATGTCATAGCTGTTCTGTGAAATGTTATCCGCTCACAATTC 1556
Db 3146 GCTTGGGCTAATCATGTCATAGCTGTTCTGTGAAATGTTATCCGCTCACAATTC 3205
Qy 1557 CACAAACATAGACGCGGAAGCATAAAGTTAAAGCTTGGGCTGCTAATGATGAGCT 1616
Db 3206 CACAAACATAGAGCGCGGAAGCATAAAGTTAAAGCTTGGGCTGCTAATGATGAGCT 3265
Qy 1617 AACTCATTAATTTGGCTGCTCACTGCGCGCTTCCAGTCGCGAAACCTGTGCTGC 1676
Db 3266 AACTCATTAATTTGGCTGCTCACTGCGCGCTTCCAGTCGCGAAACCTGTGCTGC 3325
Qy 1677 AGTGCATTAATGATCGGCAACGCGCGGAGAGAGCGGTTTGGCTATTTGGGCGCTCTT 1736
Db 3326 AGTGCATTAATGATCGGCAACGCGCGGAGAGAGCGGTTTGGCTATTTGGGCGCTCTT 3385
Qy 1737 CGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1796
Db 3386 CGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445

Qy 1797 CTGACTGAAAGCGGTTAATACGTTATCCAGATTCAGGGATTAACGAGAAAGACA 1856
Db 3446 CTGACTGAAAGCGGTTAATACGTTATCCAGATTCAGGGATTAACGAGAAAGACA 3505
Qy 1857 TGTGAGCAAAAGCCAGCAAAAGCCAGGAAACGTTAAAAAGCCGCTTGGCTTTT 1916
Db 3506 TGTGAGCAAAAGCCAGCAAAAGCCAGGAAACGTTAAAAAGCCGCTTGGCTTTT 3565
Qy 1917 TCCATAGGCTCCGCCCCCTGACAGCATTCACAAAATGACGCTCAAGTCAGAGTGC 1976
Db 3566 TCCATAGGCTCCGCCCCCTGACAGCATTCACAAAATGACGCTCAAGTCAGAGTGC 3625
Qy 1977 GAAACCCGACAGGATATTAAGATACAGAGGTTTCCCTGAAAGCTCCCTGCGCT 2036
Db 3626 GAAACCCGACAGGATATTAAGATACAGAGGTTTCCCTGAAAGCTCCCTGCGCT 3685
Qy 2037 CTCTGTTCCGACCTCGCTTACCGGATACCTGTCCGCTTCTCTGCGGAGAGC 2096
Db 3686 CTCTGTTCCGACCTCGCTTACCGGATACCTGTCCGCTTCTCTGCGGAGAGC 3745
Qy 2097 TGGCGCTTCTGATAGCTCAGCTGATGATCTGATTTGCTGATAGTCTGCTCA 2156
Db 3746 TGGCGCTTCTGATAGCTCAGCTGATGATCTGATTTGCTGATAGTCTGCTCA 3805
Qy 2157 AGCTGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTACT 2216
Db 3806 AGCTGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTACT 3865
Qy 2217 ATGCTTGAATCCAAACCGGTAACACAGCTTATGCGCATGCGACAGCCACTGCTA 2276
Db 3866 ATGCTTGAATCCAAACCGGTAACACAGCTTATGCGCATGCGACAGCCACTGCTA 3925
Qy 2277 ACAGATTAAGCAGAGCGATATGAGCGGTGCTACAGATTCCTGAAAGTGGGCTA 2336
Db 3926 ACAGATTAAGCAGAGCGATATGAGCGGTGCTACAGATTCCTGAAAGTGGGCTA 3985
Qy 2337 ACTAGGCTACATTAAGAGCAAGTATTTGATCTGCGCTCTGTAAGCCAGTTACT 2396
Db 3986 ACTAGGCTACATTAAGAGCAAGTATTTGATCTGCGCTCTGTAAGCCAGTTACT 4045
Qy 2397 TCGAAAAAGATGTTGATCTTGTATCCGCAAAACAACCGCTGTATAGCGTGT 2456
Db 4046 TCGAAAAAGATGTTGATCTTGTATCCGCAAAACAACCGCTGTATAGCGTGT 4105
Qy 2457 TTTTGTGTCAGACGCAATTAACGCGCAAAAAAGGATCTCAAGAGATCTTTGA 2516
Db 4106 TTTTGTGTCAGACGCAATTAACGCGCAAAAAAGGATCTCAAGAGATCTTTGA 4165
Qy 2517 TCTTTTCAAGGGGCTGACGCTCAGTGAACGAAACCTACGTTAAGGATTTGCTCA 2576
Db 4166 TCTTTTCAAGGGGCTGACGCTCAGTGAACGAAACCTACGTTAAGGATTTGCTCA 4225
Qy 2577 TGAATTAACAAAAGGATCTTCACTGATCTTTTAAATTAATGAAGTTTAAAT 2636
Db 4226 TGAATTAACAAAAGGATCTTCACTGATCTTTTAAATTAATGAAGTTTAAAT 4285
Qy 2637 CAATCTAAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAATTAAT 2696
Db 4286 CAATCTAAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAATTAAT 4345
Qy 2697 CACTATCTCAGGATCTGCTATTTGCTTCAATCCATGTTGCTGACCTCCGCTGCT 2756
Db 4346 CACTATCTCAGGATCTGCTATTTGCTTCAATCCATGTTGCTGACCTCCGCTGCT 4405
Qy 2757 AGTAACTACAGATCGGAGAGGCTTACATCTGAGCCCAAGTCTGCAATGATCCGCA 2816
Db 4406 AGTAACTACAGATCGGAGAGGCTTACATCTGAGCCCAAGTCTGCAATGATCCGCA 4465
Qy 2817 ACCCAGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGGAGC 2876
Db 4466 ACCCAGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGGAGC 4525

QY 2877 GCAGAGAGTCTCGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCGGAG 2936
 DB 4526 GCAGAGAGTCTCGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCGGAG 4585
 QY 2937 CTAGAGTAAGTATTCGCGCAATTAATTTGCGCAAGTGTGCGCATTTGCTACAGCA 2996
 DB 4586 CTAGAGTAAGTATTCGCGCAATTAATTTGCGCAAGTGTGCGCATTTGCTACAGCA 4645
 QY 2997 TCGTGTGTCAAGCTCTGCTTTGGTATGCTTCACTTCACTCCGGTCCCAAGCATCA 3056
 DB 4646 TCGTGTGTCAAGCTCTGCTTTGGTATGCTTCACTTCACTCCGGTCCCAAGCATCA 4705
 QY 3057 GCGAGTATCAATGATCCCGCAATGTTGTCAGAAAAAGGGTTAGCTCTTCCGCTCTCGCA 3116
 DB 4706 GCGAGTATCAATGATCCCGCAATGTTGTCAGAAAAAGGGTTAGCTCTTCCGCTCTCGCA 4765
 QY 3117 TCGTGTGTCAAGTATTCGCGCAATTAATTTGCGCAAGTGTGCGCATTTGCTACAGCA 3176
 DB 4766 TCGTGTGTCAAGTATTCGCGCAATTAATTTGCGCAAGTGTGCGCATTTGCTACAGCA 4825
 QY 3177 ATTCTCTTACTGTCTATGCTCCATCCGTAAGATGCTTTCTGTGATCTGTGATCTCAACCA 3236
 DB 4826 ATTCTCTTACTGTCTATGCTCCGTAAGATGCTTTCTGTGATCTGTGATCTCAACCA 4885
 QY 3237 AGTATCTGTGAATAGTATGCTATGCGGCGCAAGTGTCTTCCGCGCTCAATACCGG 3296
 DB 4886 AGTATCTGTGAATAGTATGCTATGCGGCGCAAGTGTCTTCCGCGCTCAATACCGG 4945
 QY 3297 ATATATACCGCGCCCATATGAGCAAACTTTAAAGTGTCTATGCTATGCAAAAGCTTCTCG 3356
 DB 4946 ATATATACCGCGCCCATATGAGCAAACTTTAAAGTGTCTATGCTATGCAAAAGCTTCTCG 5005
 QY 3357 GCGCAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCAGATGTAACCCACTCTGT 3416
 DB 5006 GCGCAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCAGATGTAACCCACTCTGT 5065
 QY 3417 CACCAACTGATCTTACGATCTTTACTTCAACAGGTTCTGCGGAGCAAAAACAG 3476
 DB 5066 CACCAACTGATCTTACGATCTTTACTTCAACAGGTTCTGCGGAGCAAAAACAG 5125
 QY 3477 GAAAGCAAAATGCGCAAAAGGAAATAGGCGCAACGCAAAATGTTGAATACATAC 3536
 DB 5126 GAAAGCAAAATGCGCAAAAGGAAATAGGCGCAACGCAAAATGTTGAATACATAC 5185
 QY 3537 TCTTCTTTTCAATATTAATTAAGCATTTATCAAGGTTATTTCTCATAGCGGATCA 3596
 DB 5186 TCTTCTTTTCAATATTAATTAAGCATTTATCAAGGTTATTTCTCATAGCGGATCA 5245
 QY 3597 TATTGAATGTAATTAAGAAAAATAAACAATAGGGTTCGCGCAATTTCCCGAAGAG 3656
 DB 5246 TATTGAATGTAATTAAGAAAAATAAACAATAGGGTTCGCGCAATTTCCCGAAGAG 5305
 QY 3657 TGCCAC 3662
 DB 5306 TGCCAC 5311

RESULT 9
 AF171227/c 3552 bp DNA circular SYN 04-FEB-2000
 LOCUS Shuttie vector pSLIRES11, complete sequence.
 DEFINITION AF171227
 ACCESSION AF171227.1 GI:5764544
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Shuttie vector pSLIRES11
 Shuttie vector pSLIRES11
 artificial sequences; vectors.
 REFERENCE
 AUTHORS
 1 (bases 1 to 3552)
 Chen, C.-M., Smith, D.M., Peters, M.A., Samson, M.E., Zitz, J.,
 Tabin, C.J., and Cepko, C.L.
 TITLE
 Production and design of more effective avian
 replication-incompetent retroviral vectors
 JOURNAL
 Dev. Biol. 214 (2), 370-384 (1999)

MEDLINE 99456779
 PUBMED 10525341
 REFERENCE 2 (bases 1 to 3552)
 AUTHORS Chen, C.-M.A., Samson, M.E.S. and Cepko, C.L.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-1999) Genetics, Harvard Medical School, 200
 Longwood Avenue, Boston, MA 02115, USA
 FEATURES
 SOURCE
 1..3552
 /organism="Shuttie vector pSLIRES11"
 /mol_type="genomic DNA"
 /db_xref="taxon:101936"
 /lab_host="Escherichia coli"
 join(2864..3552,1..2208)
 /note="similar to pBluescript"
 2236..2789
 /note="IRES of EMCV"
 2790..2849
 /note="multiple cloning site"
 ORIGIN
 Query Match 71.3%; Score 2610.8; DB 12; Length 3552;
 Best Local Similarity 83.8%; Pred. No. 0;
 Matches 3070; Conservative 0; Mismatches 482; Indels 110; Gaps 6;
 QY 1 CTAATATTGTAAGCTTAATATTTGTTAAATTTGCGTTAAATTTGTTAAATCAAGTC 60
 DB 3552 CTAATATTGTAAGCTTAATATTTGTTAAATTTGCGTTAAATTTGTTAAATCAAGTC 3493
 QY 61 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTAATCAAAAGAAATGACCGA 120
 DB 3492 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTAATCAAAAGAAATGACCGA 3433
 QY 121 GATAGGGTGAAGTGTGTTCCAGTTTGAACCAAGTCCACTATTAAGAAAGTGAAGTC 180
 DB 3432 GATAGGGTGAAGTGTGTTCCAGTTTGAACCAAGTCCACTATTAAGAAAGTGAAGTC 3373
 QY 181 CAAGCTCAAGAGGCAAAACCGCTTATCAAGGCGGATGCGCCACTAGTGAACCATCAC 240
 DB 3372 CAAGCTCAAGAGGCAAAACCGCTTATCAAGGCGGATGCGCCACTAGTGAACCATCAC 3313
 QY 241 CTAATCAAGTTTGTGGGGTGAAGTGCCTGTAAGCACTAAATCGAAACCTTAAGGAG 300
 DB 3312 CTAATCAAGTTTGTGGGGTGAAGTGCCTGTAAGCACTAAATCGAAACCTTAAGGAG 3253
 QY 301 CCCCAGATTAGACTTGAACGGGAAAGCCGCGCAACGTGCGAAGAAAGGAAAGAA 360
 DB 3252 CCCCAGATTAGACTTGAACGGGAAAGCCGCGCAACGTGCGAAGAAAGGAAAGAA 3193
 QY 361 AGCGAAAGAGCGGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAACGCTGCGCTAAC 420
 DB 3192 AGCGAAAGAGCGGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAACGCTGCGCTAAC 4133
 QY 421 CACACCGCGCGCTTAATGCGCGCTAACAGGGGCGGCTCCCATTCGCGCATTCAGGCTG 480
 DB 3132 CACACCGCGCGCTTAATGCGCGCTAACAGGGGCGGCTCCCATTCGCGCATTCAGGCTG 3073
 QY 481 CAACTGTTGGAAAGGCGATCGTGCAGGCTCTTCTGCTAATTAAGCGCAAGTGCAGAAAG 540
 DB 3072 CAACTGTTGGAAAGGCGATCGTGCAGGCTCTTCTGCTAATTAAGCGCAAGTGCAGAAAG 3013
 QY 541 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAAGCCAGGGTTTCCAGTCAAGAGCTTG 600
 DB 3012 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAAGCCAGGGTTTCCAGTCAAGAGCTTG 2953
 QY 601 TAAACGAGCGGCGAGTGAAGCGCGCTGTCATTCAGTTTGAACCCGAGGAGAGCG 660
 DB 2952 TAAACGAGCGGCGAGTGAAGCGCGCTGTCATTCAGTTTGAACCCGAGGAGAGCG 2913
 QY 661 GCAGACTCGCGGTCAATGTGTTTACAGGCTGATGAGAGCAATGAGTCTGACAC 720
 DB 2912 GCAGACTCGCGGTCAATGTGTTTACAGGCTGATGAGAGCAATGAGTCTGAGGCT 2853

Db 782 AAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTGTCGGGAAGCTAG 723
Qy 2941 AGTAAGTATGTTCCGCACTTAATAGTTTGGCGAAGCTTGTGCCATGTCTACAGAGCATGT 3000
Db 722 AGTAAGTATGTTCCGCACTTAATAGTTTGGCGAAGCTTGTGCCATGTCTACAGAGCATGT 663
Qy 3001 GGTCGACGCTGCTGTTGGTAATGCTTCACTTCAGCTCCGCTTCCCAACATCAAGGCG 3060
Db 662 GGTCGACGCTGCTGTTGGTAATGCTTCACTTCAGCTCCGCTTCCCAACATCAAGGCG 603
Qy 3061 AGTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGCTCTTCCGCTCCGATGCT 3120
Db 602 AGTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGCTCTTCCGCTCCGATGCT 543
Qy 3121 TGTCAAGATGATGTTGGCGGAGTGTATCACTCATGTTATGCGACACTGCAATATTC 3180
Db 542 TGTCAAGATGATGTTGGCGGAGTGTATCACTCATGTTATGCGACACTGCAATATTC 483
Qy 3181 TCTTACGTGTCATGCGCATCCGTAAGATGCTTTTCTGTGCTGTGATCTCAACCAAGTC 3240
Db 482 TCTTACGTGTCATGCGCATCCGTAAGATGCTTTTCTGTGCTGTGATCTCAACCAAGTC 423
Qy 3241 ATTCTGAATATGATGTCGCGGCGACCGAGTGTCTCTTCCGCGCTCAATACGGATTA 3300
Db 422 ATTCTGAATATGATGTCGCGGCGACCGAGTGTCTCTTCCGCGCTCAATACGGATTA 363
Qy 3301 TACCGCGCCACATAGCAGAACTTTAAAGTCTCATCTTGAAGAAAGCTTCTTGGGGCG 3360
Db 362 TACCGCGCCACATAGCAGAACTTTAAAGTCTCATCTTGAAGAAAGCTTCTTGGGGCG 303
Qy 3361 AAAAATCTGCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCCGTCAC 3420
Db 302 AAAAATCTGCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCCGTCAC 243
Qy 3421 CAATGATCTTCAAGATCTTCTTACCTTCAAGAGCTTCTGAGTGAAGAAAGAGAG 3480
Db 242 CAATGATCTTCAAGATCTTCTTACCTTCAAGAGCTTCTGAGTGAAGAAAGAGAG 183
Qy 3481 GCAAAATGCGCAAAAAGGAATGAGGCGACACGGAATGTTGAATATCTCATCTT 3540
Db 182 GCAAAATGCGCAAAAAGGAATGAGGCGACACGGAATGTTGAATATCTCATCTT 123
Qy 3541 CCTTTTCAATATATGAAACATTTATCAAGGTTATGTTCTCATGAGCGATCATATT 3600
Db 122 CCTTTTCAATATATGAAACATTTATCAAGGTTATGTTCTCATGAGCGATCATATT 63
Qy 3601 TGAATGATATTAAGAAAAATAAACAATAGGGGTTCCGCGCACTTTCCCGAAAAAGTCC 3660
Db 62 TGAATGATATTAAGAAAAATAAACAATAGGGGTTCCGCGCACTTTCCCGAAAAAGTCC 3
Qy 3661 AC 3662
Db 2 AC 1

RESULT 10
AX306330 4001 bp DNA linear PAT 11-DEC-2001
LOCUS Sequence 13 from Patent WO018121.
DEFINITION AX306330
ACCESSION AX306330
VERSION AX306330.1 GI:17645569
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1 Plaetnick, G., Renard, J.P. and Bogaert, T.
AUTHORS
TITLE Vector constructs
JOURNAL Patent: WO 018121-A 13 22-NOV-2001;
Devgen NV (BE)
FEATURES
source location/Qualifiers
1..4001
/organism="synthetic construct"

ORIGIN
Query Match 71.3%; Score 2609.4; DB 6; Length 4001;
Best Local Similarity 83.1%; Pred. No. 0;
Matches 3069; Conservative 0; Mismatches 591; Indels 35; Gaps 7;
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pGN59A"

1 CTAATTTGAAGCGTTAATATTTTGTAAATTCGCTTAAATTTTGTAAATCAGCTC 60
Db 63 CGAATTTGAAGCGTTAATATTTTGTAAATTCGCTTAAATTTTGTAAATCAGCTC 122
Qy 61 ATTTTGAAGCGTTAATATTTTGTAAATTCGCTTAAATTTTGTAAATCAGCTC 120
Db 123 ATTTTGAAGCGTTAATATTTTGTAAATTCGCTTAAATTTTGTAAATCAGCTC 182
Qy 121 GATAGGGTGAAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGAAGT 180
Db 183 GATAGGGTGAAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGAAGT 242
Qy 181 CAAGTCAAGAGGCGGAAACCGCTCTATCAGGGGATGCGCCACTAGTGAACCATCAC 240
Db 243 CAAGTCAAGAGGCGGAAACCGCTCTATCAGGGGATGCGCCACTAGTGAACCATCAC 302
Qy 241 CTAATCAAGTTTGTGGGGTGAAGTCCGTTAAAGCACTAATGGAACCTTAAAGGAG 300
Db 303 CAATCAAGTTTGTGGGGTGAAGTCCGTTAAAGCACTAATGGAACCTTAAAGGAG 362
Qy 301 CCGCGGATTAAGCTTGAAGCGGGAAGCGGCGGCAAGTGGCGAAGAAAGGGAAGAA 360
Db 363 CCGCGGATTAAGCTTGAAGCGGGAAGCGGCGGCAAGTGGCGAAGAAAGGGAAGAA 422
Qy 361 AGCGAAAGAGCGGCGCTAGGCGCTGGAAGTGAAGCGGCTCAAGCTGCGCTTACAC 420
Db 423 AGCGAAAGAGCGGCGCTAGGCGCTGGAAGTGAAGCGGCTCAAGCTGCGCTTACAC 482
Qy 421 CACACCGCGCGCTTAAATGCGCGCTTAAAGCGGCTCCATTCGCTTCAAGCTGCG 480
Db 483 CACACCGCGCGCTTAAATGCGCGCTTAAAGCGGCTCCATTCGCTTCAAGCTGCG 541
Qy 481 CAATGTTGGAAGGAGTGGGCGGCTGTTGCTTAAAGCGGCTGCGGGAAGG 540
Db 542 CAATGTTGGAAGGAGTGGGCGGCTGTTGCTTAAAGCGGCTGCGGGAAGG 601
Qy 541 GGGATGTCGCAAGCGATTAAGTGGTAAAGCGGCTTCCAGTCAAGAGCTG 600
Db 602 GGGATGTCGCAAGCGATTAAGTGGTAAAGCGGCTTCCAGTCAAGAGCTG 661
Qy 601 TAAAGCAAGCGGCGAGTGAAGCGGCTT-----CGTTCAATTCAGCTTTTGA 646
Db 662 TAAAGCAAGCGGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 721
Qy 647 CCGGTGAAGAGCGGCGAGTCCGCGGTCGCAAGTGTGTTTAAAGCGTGAAGAGAGT 706
Db 722 CCGGTGAAGAGCGGCGAGTCCGCGGTCGCAAGTGTGTTTAAAGCGTGAAGAGAGT 781
Qy 707 AAGATGCTGCAAGCGTGAAGAGCGGCTGAAGT-----TAACTTGAAGAGTAA 759
Db 782 CCGGTGAAGAGCGGCGAGTGAAGAGCGGCTGAAGT-----TAACTTGAAGAGTAA 841
Qy 760 TCAATATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 812
Db 842 GAATATATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 901
Qy 813 CTGTAATAGAGTCAATATGAGGGAATGAGGGAATGAGGGAATGAGGGAATGAGGGA 872
Db 902 TATTAAGAGTCAATATGAGGGAATGAGGGAATGAGGGAATGAGGGAATGAGGGAAT 961
Qy 873 ATGATAGAGTCAATATGAGGGAATGAGGGAATGAGGGAATGAGGGAATGAGGGAAT 932
Db 962 ATGATAGAGTCAATATGAGGGAATGAGGGAATGAGGGAATGAGGGAATGAGGGAAT 1021

QY 933 GCCACGCGGTGAGCTCCAGCTTTTGTCCCTTATGAGAGGTTAATTAGATCCATGC 992
Db 1022 TGGAAAGCCGAAATAGAGAGAGATGCTGAGGTGCGCCGTTTATTGAATGAAGCC 1081
QY 993 GTCAATTTTACGAGACTA--TCTTTAGAGGTTAATCTGCTGCATCAGATCATATC 1049
Db 1082 TCTTTTGTGACGAGAAACAGGACTGGAATGACGTTTAAAGTTTACCTATTAAG 1141
QY 1050 GTGCGGTCCTTTTCCGCTCAATCATGCCCCAACTGCGCTATCGGCGATCGGAG 1109
Db 1142 AGAGAGCCGTATCTCTGTTTGTGAGATGACAGAGATATTAATGACAGCCGCGG 1201
QY 1110 GAAGAAGCCCGTCCCTTTTCCGCGAGTTGAAAGCGGCATGGAAGGTTTGCAGAT 1169
Db 1202 ACGATGTGATCCCGCTGCGC-AGTGACGTCTCTTAAGCGTAAAGTCCCGTGAAC 1260
QY 1170 GACTGCTGCTG-CATTGACGTTGAGGAAACGACGTTTACATGATGATTCGGAAG 1227
Db 1261 TTTTACCGGTGTGATCATATCGGAGTGAAGCTGCGCATGATGACACCGATATGCGCA 1320
QY 1228 GTTGGCCATGACGCGCTTTTAAAGGTGAATCTGTTGATGAGCGCATCTGGATACCAATT 1287
Db 1321 GTTGGCCGTCCTCGTTATCGGAGAGAAAGTGTCTATCTACGCAACCGCAAAATGACA 1380
QY 1288 CGTGGCGCTTTTCCGACACAGTTCGGAATGTGACGCCGAAGCGCATCAGCAACCGCA 1347
Db 1381 TCAGAAACGCCATTTAACTGATGTTCTGCGGAATTAATATGACGCTCCCTTATACACA 1440
QY 1348 ACAATACCGGCGACAGCCGGAATCTGCGCGGTGTGACATTAATGACAGCGGTGCG 1407
Db 1441 GCGTTTCAAGCAATATGATCTGAGAGGATCTTCCATACCTACCGATCTGCGCTGAG 1500
QY 1408 CGTGGGATTTATCGTACAGCGAGAGGATCTGCGTGTGATCGCGCAAAATGACAT 1467
Db 1501 GTGCGGCGCGAATCTCTAGAGTGAAAGCTTCTCGCCATATGATGATGATTTACAG 1560
QY 1468 GGATACCCCGTATTAACCGGCGGCGCGCTGCGGATCAATGATGATGATGATGATGAT 1527
Db 1561 CTGAGTATTTATAGTGTCACTTAATAGCTTGAATGATGATGATGATGATGATGATGAT 1620
QY 1528 TGTGGAATTTGTTTCCGCTCAATTCACACAAATACAGAGCGGAGCAATTAAGT 1587
Db 1621 TGTGGAATTTGTTTCCGCTCAATTCACACAAATACAGAGCGGAGCAATTAAGT 1680
QY 1588 TAAAGCTGAGGAGTCTATAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1647
Db 1681 TAAAGCTGAGGAGTCTATAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1700
QY 1648 CGCTTCCAGTCCGGAACCTGTGTGCACTGATTAATGATGATGATGATGATGATGATGATGAT 1740
Db 1741 CGCTTCCAGTCCGGAACCTGTGTGCACTGATTAATGATGATGATGATGATGATGATGATGAT 1800
QY 1708 GAGAGCGGTTTGTGATTTGAGGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 1767
Db 1801 GAGAGCGGTTTGTGATTTGAGGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 1860
QY 1768 GGTGTTCCGCTGCGGAGCGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1827
Db 1861 GGTGTTCCGCTGCGGAGCGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
QY 1828 AGATATGAGGAGTAAAGCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1887
Db 1921 AGATATGAGGAGTAAAGCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1888 CGTAAAGAGCGGCTTGTGCTGCTGCTTTCATAGGCTCGCCCTCGACGAGATCA 1947
Db 1981 CGTAAAGAGCGGCTTGTGCTGCTGCTTTCATAGGCTCGCCCTCGACGAGATCA 2040
QY 1948 CAAAAATGACGCTCAAGTCAAGGTGAGGAAACCGGACAGATATTAAGATCAAGG 2007
Db 2041 CAAAAATGACGCTCAAGTCAAGGTGAGGAAACCGGACAGATATTAAGATCAAGG 2100
QY 2008 GTTTCCTGGAAGCTCCCTGCTGCTCTCTGTTCCAGCTGCGGCTTACCGGAT 2067

Db 2101 GTTTCCTGGAAGCTCCCTGCTGCTCTCTGTTCCAGCTGCGGCTTACCGGAT 2160
QY 2068 CCGTCCGCTTTCCTTCCGAGAGCGTGGCGCTTTCATAGTCAAGCTGATGAT 2127
Db 2161 CCGTCCGCTTTCCTTCCGAGAGCGTGGCGCTTTCATAGTCAAGCTGATGAT 2220
QY 2128 TCTGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2187
Db 2221 TCTGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
QY 2188 GCCGACCGGTCGCTTATCCGCTTAACTATGCTCTTGAAGTCAACCGGATGACAGCA 2247
Db 2281 GCCGACCGGTCGCTTATCCGCTTAACTATGCTCTTGAAGTCAACCGGATGACAGCA 2340
QY 2248 CTTATGCGCATGCGAGCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2307
Db 2341 CTTATGCGCATGCGAGCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2400
QY 2308 TGTATGAGGTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2367
Db 2401 TGTATGAGGTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
QY 2368 TATCTGCGCTTGTGAGAGCAGTTCCTTGGAAAAAGTGTGATGATGATGATGATGAT 2427
Db 2461 TATCTGCGCTTGTGAGAGCAGTTCCTTGGAAAAAGTGTGATGATGATGATGATGAT 2520
QY 2428 CAAACAAACCGCTGTGATGCGGTTTGTGTAAGACAGATTAACGCGAG 2487
Db 2521 CAAACAAACCGCTGTGATGCGGTTTGTGTAAGACAGATTAACGCGAG 2580
QY 2488 AAAAAAGATCTGAAGAGATCTTGTGATCTTTCAGCGGCTGACGCTCATGAGAA 2547
Db 2581 AAAAAAGATCTGAAGAGATCTTGTGATCTTTCAGCGGCTGACGCTCATGAGAA 2640
QY 2548 GGAAGATCTGAAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2607
Db 2641 GGAAGATCTGAAGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 2608 CTTTAAATTTAAATGAATTTTAAATCAATCTAATATATATATATATATATATATATATAT 2667
Db 2701 CTTTAAATTTAAATGAATTTTAAATCAATCTAATATATATATATATATATATATATATAT 2760
QY 2668 TGAAGTATCAATGCTTATATGATGAGCAGCTATCTGAGGATCTGATTTTCGTT 2727
Db 2761 TGAAGTATCAATGCTTATATGATGAGCAGCTATCTGAGGATCTGATTTTCGTT 2820
QY 2728 ATCCATAGTTGCTGATCTCCCGTGTGATGATTAATCAATGATGATGATGATGATGAT 2787
Db 2821 ATCCATAGTTGCTGATCTCCCGTGTGATGATTAATCAATGATGATGATGATGATGAT 2880
QY 2788 TGGCCCCAGTGTGATGATATACCGGAGACCAAGCTCACCGGCTCAAGATTTATACAG 2847
Db 2881 TGGCCCCAGTGTGATGATATACCGGAGACCAAGCTCACCGGCTCAAGATTTATACAG 2940
QY 2848 AATTAACAGCAGCGGAGGAGGAGCGAGCGAGAGGATGATGATGATGATGATGATGAT 2907
Db 2941 AATTAACAGCAGCGGAGGAGGAGCGAGCGAGAGGATGATGATGATGATGATGATGAT 3000
QY 2908 CATCAAGTCTATTAATTTTTCGCGGAAAGCTAGATGATGATGATGATGATGATGATGAT 2967
Db 3001 CATCAAGTCTATTAATTTTTCGCGGAAAGCTAGATGATGATGATGATGATGATGATGAT 3060
QY 2968 GCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3027
Db 3061 GCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
QY 3028 TTTCAATGCTCGGTTTCCAGAGATCAAGCGAGTTCATGATGATGATGATGATGATGAT 3087
Db 3121 TTTCAATGCTCGGTTTCCAGAGATCAAGCGAGTTCATGATGATGATGATGATGATGAT 3180
QY 3088 AAAAGCGTTAGCTCTTGGGCTCTCCGATCTGTTGCAAGATGATGATGATGATGATGAT 3147

Db 3181 AAAAGCGTAGTCTCTCGGTCCTCCGATCGTGTGTCAGAAATGTAAGTTGGCCGAGTGT 3240
Qy 3148 ATCATCTAGTGTATGGACGACATGCAATATTTCTTACTGTATGTCATGCCATCCGTAAAGTG 3207
Db 3241 ATCACTAATGGTATGGACGACATGCAATATTTCTTACTGTATGTCATGCCATCCGTAAAGTG 3300
Qy 3208 CTTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTTGAGAAATGATGTATGCGCGGAC 3267
Db 3301 CTTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTTGAGAAATGATGTATGCGCGGAC 3360
Qy 3368 GAGTTGCTCTTCCCGCGGCAATACGGGATTAATACCGGCACATAGCAGAACTTTTAA 3327
Db 3361 GAGTTGCTCTTCCCGCGGCAATACGGGATTAATAGTGTATAGACATAGCAGAACTTTTAA 3420
Qy 3328 AGTGTCTCATCTTGGAAAAAGTTCTTGGGGCGAAAACTTCACAGATCTTACCGCTGT 3387
Db 3421 AGTGTCTCATCTTGGAAAAAGTTCTTGGGGCGAAAACTTCACAGATCTTACCGCTGT 3480
Qy 3388 GAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACTT 3447
Db 3481 GAGATCCAGTTCGATGTAAACCACTCGTGCACCCCACTGATCTTCAGCATCTTTTACTT 3540
Qy 3448 CACCAAGCGTTTCTGGGTGAGCAAAAACAGAAAGCAAAATGCCGCAAAAAAGGAAATAG 3507
Db 3541 CACCAAGCGTTTCTGGGTGAGCAAAAACAGAAAGCAAAATGCCGCAAAAAAGGAAATAG 3600
Qy 3508 GGGCAGCAGGAAATGTTGAATATCTATCTCTCTTTTCAATATTTTGAAGCATTTA 3567
Db 3501 GGGCAGCAGGAAATGTTGAATATCTATCTCTCTTTTCAATATTTTGAAGCATTTA 3660
Qy 3568 TCAGGGTATATGTCTCATGAGCGGATACATATTTGAATGTAATTTGAAGAAATTAACAAT 3627
Db 3661 TCAGGGTATATGTCTCATGAGCGGATACATATTTGAATGTAATTTGAAGAAATTAACAAT 3720
Qy 3628 AGGGGTTCCGCGCACATTTCCCGAAAAAGTCCAC 3662
Db 3721 AGGGGTTCCGCGCACATTTCCCGAAAAAGTCCAC 3755

RESULT 11
AF179627 5540 bp DNA linear SYN 22-NOV-1999
LOCUS AF179627
DEFINITION Expression vector pCENEX645, complete sequence.
ACCESSION AF179627
VERSION AF179627.1 GI:6457303
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Bugster,H.P., Bartelch,S., Wurgler,F.B. and Bengtstag,C.
TITLE Functional co-expression of human oxidoreductase and cytochrome P450 1A1 in Saccharomyces cerevisiae results in increased EROD activity
JOURNAL Biochem. Biophys. Res. Commun. 185 (2), 641-647 (1992)
MEDLINE 92304288
PubMed 1610357
REFERENCE
AUTHORS Bengtstag,C. and Paladino,G.
TITLE A collection of transformed Saccharomyces cerevisiae strains as an unlimited source of human enzymes
JOURNAL Bioworld 97 (6), 2-7 (1997)
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (24-AUG-1999) Genetics, Institute of Toxicology, ETH Zurich, Schorenstrasse 16, Schwerzenbach 8603, Switzerland
FEATURES
Source
1..5540
/organism="Expression vector pCENEX645"
/mol_type="genomic DNA"
/db_xref="taxon:106613"
/note="yeast expression vector pCENEX645 (alias pCS645)"

containing truncated GAPDH promoter for constitutive expression of cDNAs in yeast"
417..1220
/gene="URA3"
417..1220
/gene="URA3"
/note="yeast selection marker"
/codon_start=1
/product="URA3"
/protein_id="AA09466.1"
/db_xref="GI:6457304"
/translation="MSKATYKERATPSPVAAKLFINIHEKOTNLCAIDVPTTKEL
LEIVLALGPICILKTHVDLTDFSMGTCPLKALSAKXNFLLFEORKEFADICNTYK
LOYSAGVYRIAEWADITTAAGVUGRGVSCVAKOAEVTEKREGLTLALRLSCGSL
TGEYTKGTVDIAKSDKDFVIGTIAQRDMGDESDYDLINTPGGLDKCTALQQR
TVDDVSTGSDIILVGRGLAKGDAVBERRYRKAGWEAVLRRCGQN"
2310..2508
/note="truncated GAPDH promoter"
complement (3958..4818)
/gene="Amp"
complement (3958..4818)
/gene="Amp"
/note="ampicillin resistance; Escherichia coli selection marker"
/codon_start=1
/product="Amp"
/protein_id="AA09467.1"
/db_xref="GI:6457305"
/translation="MSIOHFRVALIPEFAFCLPVFAHFDLVFKDCEDQLGARVY
LEIDNSGKILIESFRPEPFPMSFTFVLGAVLFDIDGOEOLGRIRIHSQNDLYE
YSPVTEKHLIDGMVRELGSAAITMSNTANILITIGPKELTAFLHMGDVTRL
DRWPELNEAIPNDERDTTPVMAATTLRLKLTSELTLASROOLIMWEDKVAQPL
LRSLPFWMTIADSKSGERSKGIINALGPDGKPSRIVVIYTGSOATMDERNRQIA
EIGASLIKHW"

ORIGIN
Query Match 69.9%; Score 2558.8; DB 12; Length 5540;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 3035; Conservative 0; Mismatches 557; Indels 72; Gaps 7;
Qy 3 AAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 62
Db 1351 AAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 1410
Qy 63 TTTTAAACCAATAGCGCGAAATTCGCAAAATCCCTTATTAATCAAAAGATAGACCGAGA 122
Db 1411 TTTTAAACCAATAGCGCGAAATTCGCAAAATCCCTTATTAATCAAAAGATAGACCGAGA 1470
Qy 123 TAGGGTGAAGTGTTCGAGTTTGGAAACAAGTCCACTATTAAGAAGGTGAGCTCCA 182
Db 1471 TAGGGTGAAGTGTTCGAGTTTGGAAACAAGTCCACTATTAAGAAGGTGAGCTCCA 1530
Qy 183 ACGTCAAAAGGCGAAAAACGCTCATAGGGCGATGCGCACTACGTGAACCATCACCT 242
Db 1531 ACGTCAAAAGGCGAAAAACGCTCATAGGGCGATGCGCACTACGTGAACCATCACCT 1590
Qy 243 AATCAAGTTTTCGGGTCGAGGTCCGCTAAGACATTAATCGAACCTTAAGGGAGCC 302
Db 1591 AATCAAGTTTTCGGGTCGAGGTCCGCTAAGACATTAATCGAACCTTAAGGGAGCC 1650
Qy 303 CCGCATTTAAGCTTGAACGGGGAAGCGCGCAACGTGGGGAAGAAGGAGGAAGAAG 362
Db 1651 CCGCATTTAAGCTTGAACGGGGAAGCGCGCAACGTGGGGAAGAAGGAGGAAGAAG 1710
Qy 363 CGAAAGAGCGGCGCTAGGGCGCTGCAAGTGAAGCGTCAACGCTGCGCTAAGCAACA 422
Db 1711 CGAAAGAGCGGCGCTAGGGCGCTGCAAGTGAAGCGTCAACGCTGCGCTAAGCAACA 1770
Qy 423 CACCGCGCGCTTAATGCGCGCTACAGGCGCGTGC---CATTCGCCATTACGGCTGC 479
Db 1771 CACCGCGCGCTTAATGCGCGCTACAGGCGCGTGC---CATTCGCCATTACGGCTGC 1830
Qy 480 GCAACTGTGGGAAGGCGAATCGGTGCGGCGCTTCTGCTATTACGCACTGCGCAAG 539

Db 1831 GCACTGTTGGAAAGGCGATCGGTGCGGCTCTTCGCTATTAGCCAGCTGGCGAAAG 1890
Qy 540 GGGATGTGCTGCAAGGCGATTAAAGTGGGTAAACGACAGGTTTTCCAGTCAAGACGTT 599
Db 1891 GGGGATGTGCTGCAAGGCGATTAAAGTGGGTAAACGACAGGTTTTCCAGTCAAGACGTT 1950
Qy 600 GTAAACGACGCGCAGTGAAGCGCGCTCGTTCAATTCACGTTTTTAAACCCGAGAGACG 659
Db 1951 GTAAACGACGCGCAGTGA-----ATTGTAAATACGACTCACTAATAGGCCAATTTGG 2001
Qy 660 GGCAGACTCGCGGTGCMAATGTGTTTTACAGGTGATAGACAGATGAAGATGCTGCACA 719
Db 2002 AGCTCCACCGCGGTG-----GCGCGCTCTTAAGTCAGCGCTCTCC 2042
Qy 720 CGCTGCAGAACGCGAGTGAATTAACTCTGAAGAAAGATTAATATATTGTGACGTAGTTA 779
Db 2043 TTATGCACTCTGCAATTAGGAAGCAGCTCAGTAGTGAAGCGGTTGAGCGGTTGACACCGCC 2102
Qy 780 AAGATTAATCATCGTAAATTTGACGCAATGGGATCTGTAAATACACTATATAGGCGAA 839
Db 2103 GCGCGAAGAAATGGTGCAATGCAAGAAATGGCGCCACAGTCCCCGCGCACGCGGCGCT 2162
Qy 840 TTGGTACCGGAGCCCCCTCGAGGTGACGCGTATGCAATGCAATGCAATTCGATTCG 899
Db 2163 GCCACATACCAAGCGCGGAAACAGCGCTCATGAGCGCGAAATGGGACCGGATCTTCC 2222
Qy 900 CAGCCCGGGGATCACTAGTTCTAGAGCGCGCGCACGCGGTGAGCTCCAGCTTTTG 959
Db 2223 CCAATGGTGAATGTCGGGATATAGCGCGCACCGCACCTGTGGCGCGGTGATGCGG 2282
Qy 960 TTCCCTTAAATGAGGGTTAATTAGATCCATGCGTCAATTTTACGACAGTATCTTTCTA 1019
Db 2283 GCCACAGATGGTCCGGGTGAGAGATCTCTGMAAAAGGTGAACCAAGTTCCTG 2342
Qy 1020 GGGTAACTAGCTGATAGATCAATATCGTGGGTCTTTTCCGGCTCAATGATCGC 1079
Db 2343 AAATTAATTC-----CCTTACTGACTAATTAAGTAAATTA 2376
Qy 1080 CCAAGCTGCGCTATCTGGGATCGGGAGGAAAGCCGCTGCTTTTCCCGAGAGTT 1139
Db 2377 AGACGGTAGATTAATGATTTCTGTA-----AATCTATTTCTTAACTTTCTTAATT 2431
Qy 1140 GAAGCGGATGMAAGAGTTTGCAGAGATGACTGCTGCAATGACGTTGACGAA 1199
Db 2432 CTACTTTTAATGATGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2482
Qy 1200 CGCAGCTTACCATGATGATTGGGAAGGTGCGCATGCAAGCTTTTAAAGGTGAATG 1259
Db 2483 TTGCAATTAACACATTAATTAAGAAATTAATGATTTGCAAGCTTTCTTTTCTTTTCTTTT 2542
Qy 1260 TTGCTGAGCGCACCTGGGATACCAATTCGTGCGCGCTTTTCCGACACAGATCCGAGT 1319
Db 2543 GCTGTTTTGCAAGCAAAATATCTGCAATCTTAGCGTCTGAGGGAACATTAATTAATG 2602
Qy 1320 GTGAGCCGGAAGCGCATCAGCAACCGGAATTAACGCGCAAGCGCGAATGCGCGTCC 1379
Db 2603 AACCCAGTCTAATTTGAGACAAATGTTTTGTAATTAATTAATTAATTAATTAATTAATG 2662
Qy 1380 GGTTGCAATTAATTAACAGCGGTGCGCGCTGGAATTAATTAATTAATTAATTAATTAATG 1439
Db 2663 AATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2722
Qy 1440 CTTGCTGATGCGCAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 2782
Db 2723 CGAGGGGGGGCGCGCGCGCGCGCGCTTTGTTCTCTTTTAAATTAATTAATTTAAATTTAAATTT 2842
Qy 1499 TTGGCTATCATAGTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTG 1558
Db 2783 TTGGCTATCATAGTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTG 2842
Qy 1559 CACAACTATGAGCGGAGGATTAAGTTAAACCTGGGGTGTAAATGATGAGCTAA 1618

Db 2843 CACAACTATGAGCGGAGGATTAAGTTAAACCTGGGGTGTCTAATGATGAGCTAA 2902
Qy 1619 CTCACATTAATTCGTTGCGCTCACTGCGGCTTTCCAGTCGGAAACCTGTGTCAG 1678
Db 2903 CTCACATTAATTCGTTGCGCTCACTGCGGCTTTCCAGTCGGAAACCTGTGTCAG 2962
Qy 1679 CTCATTAATTAATTCGCGCAACGCGCGGAGAGCGGTTGCGTATGCGGCTCTTCC 1738
Db 2963 CTCATTAATTAATTCGCGCAACGCGCGGAGAGCGGTTGCGTATGCGGCTCTTCC 3022
Qy 1739 GCTTCTGCTCACTGACTGCTGCGCTGCGTCTGCGTCTGCGGAGCGGTATCACT 1798
Db 3023 GCTTCTGCTCACTGACTGCTGCGCTGCGTCTGCGTCTGCGGAGCGGTATCACT 3082
Qy 1799 CACTCAAAAGCGGTATTAAGCTTTATCAAGAAATGAGGATTAACGAGAAAGAAATG 1858
Db 3083 CACTCAAAAGCGGTATTAAGCTTTATCAAGAAATGAGGATTAACGAGAAAGAAATG 3142
Qy 1859 TGAGCAAAAGCGCAGCAAAAGCCAGAAACGTTAAAGCGCGGTTGCTGCGTTTTC 1918
Db 3143 TGAGCAAAAGCGCAGCAAAAGCCAGAAACGTTAAAGCGCGGTTGCTGCGTTTTC 3202
Qy 1919 CATAGGCTCGCGCCCCCTGAGCAGATCAAAATGCAAGCTCAAGAGTGGCGA 1978
Db 3203 CATAGGCTCGCGCCCCCTGAGCAGATCAAAATGCAAGCTCAAGAGTGGCGA 3262
Qy 1979 AACCGCAAGACTAATTAAGATACAGAGCGTTTCCCGTGAAGCTCCCTGCGCTCT 2038
Db 3263 AACCGCAAGACTAATTAAGATACAGAGCGTTTCCCGTGAAGCTCCCTGCGCTCT 3322
Qy 2039 CCTGTTCCGACCGCGGCTTACCGGATACCTGTGCGCTTCTCCCTGCGGAGCGTG 2098
Db 3323 CCTGTTCCGACCGCGGCTTACCGGATACCTGTGCGCTTCTCCCTGCGGAGCGTG 3382
Qy 2099 GCGCTTTCTATAGCTCAAGCTGATGATCTCAGTTGCGTGAAGTCTTTGCTCCAG 2158
Db 3383 GCGCTTTCTATAGCTCAAGCTGATGATCTCAGTTGCGTGAAGTCTTTGCTCCAG 3442
Qy 2159 CTGGGCTGTGTCAGAAACCGCGCTTACCGCGACCGCTGCGCTTATCCGTAATCT 2218
Db 3443 CTGGGCTGTGTCAGAAACCGCGCTTACCGCGACCGCTGCGCTTATCCGTAATCT 3502
Qy 2219 CGTCTTAATGTCACCGCGTGAAGACAGACTTATCGCATGCGGACCGCATCTGTAC 2278
Db 3503 CGTCTTAATGTCACCGCGTGAAGACAGACTTATCGCATGCGGACCGCATCTGTAC 3562
Qy 2279 AGGATTAAGCAGAGCGGATTAAGCGGCTGCTACAGATTTCTTGAATGGTGGCTTAC 2338
Db 3563 AGGATTAAGCAGAGCGGATTAAGCGGCTGCTACAGATTTCTTGAATGGTGGCTTAC 3622
Qy 2339 TACGGCTACATGAAGACAGATTTTGTATCTGCGCTCTGCTGAAGCGCATTAATC 2398
Db 3623 TACGGCTACATGAAGACAGATTTTGTATCTGCGCTCTGCTGAAGCGCATTAATC 3682
Qy 2399 GGAATAAGTGTGATGCTTTGATCCGCGCAACAAACAGCTGTGTAGCGGTGTTT 2458
Db 3683 GGAATAAGTGTGATGCTTTGATCCGCGCAACAAACAGCTGTGTAGCGGTGTTT 3742
Qy 2459 TTTGTTTCCAGCAGCAGATTAAGCGCAGAAATAAGATCTCAAGAAATCTTTGATC 2518
Db 3743 TTTGTTTCCAGCAGCAGATTAAGCGCAGAAATAAGATCTCAAGAAATCTTTGATC 3802
Qy 2519 TTTTCTAGCGGGTGTGAGCTCAGTGGAAACGAAATCAAGTTAAGGATTTTGTCTAG 2578
Db 3803 TTTTCTAGCGGGTGTGAGCTCAGTGGAAACGAAATCAAGTTAAGGATTTTGTCTAG 3862
Qy 2579 AGATTAATCAAAAGATCTTCACTAGATCTTTTAAATTAATTAATTAATTAATTAATCA 2638
Db 3863 AGATTAATCAAAAGATCTTCACTAGATCTTTTAAATTAATTAATTAATTAATTAATCA 3922
Qy 2639 ATCTAAAGTATATATAGTAACTTGTGACAGTTACCAATGCTTAATCAGTGAAGCA 2698
Db 3923 ATCTAAAGTATATATAGTAACTTGTGACAGTTACCAATGCTTAATCAGTGAAGCA 3982

OY	2699	CCTATCTCAGGAGATCTGTCATAATTCCTGTCATCCATAGTTGGCTGACCTCCCGTGTGAG	2758
Dp	3983	CCATCTCAGGAGATCTGTCATAATTCCTGTCATCCATAGTTGGCTGACCTCCCGTGTGAG	4042
OY	2759	ATTAACTAAGATACGGAGAGGGCTTACCACTCTGGCCCCAGTGTGCATATGATACCGCGAGAC	2818
Dp	4043	ATTAACTAAGATACGGAGAGGGCTTACCACTCTGGCCCCAGTGTGCATATGATACCGCGAGAC	4102
OY	2819	CCACGGCTCACCGGCTCCAGATTTTATCAGCAATTAACCAAGCCAGCGGAGAGGGCCGAGCGC	2878
Dp	4103	CCACGGCTCACCGGCTCCAGATTTTATCAGCAATTAACCAAGCCAGCGGAGAGGGCCGAGCGC	4162
OY	2879	AGAGTGGTCCCTGCAACTTTTATCCGGCTCCATCCAGTCTATTAATTTGTTGGCCGGAGACT	2938
Dp	4163	AGAGTGGTCCCTGCAACTTTTATCCGGCTCCATCCAGTCTATTAATTTGTTGGCCGGAGACT	4222
OY	2939	AGAGTAAGTATGTTCCGCAATTAAATAGTTTGGCGACGTTGTTGCCATTGTCTACAGGCAATC	2998
Dp	4223	AGAGTAAGTATGTTCCGCAATTAAATAGTTTGGCGACGTTGTTGCCATTGTCTACAGGCAATC	4282
OY	2999	GTGGTGTACGCTCGTCTGTTTGTATGAGCTTCATTCACTCCGCTCCGATCCCAAGATCAAG	3058
Dp	4283	GTGGTGTACGCTCGTCTGTTTGTATGAGCTTCATTCACTCCGCTCCGATCCCAAGATCAAG	4342
OY	3059	CGAGTTACATGATCCCCATGTTGTGCAAAAAAGGGTTAGCTCTTTGGTCTCCGATC	3118
Dp	4343	CGAGTTACATGATCCCCATGTTGTGCAAAAAAGGGTTAGCTCTTTGGTCTCCGATC	4402
OY	3119	GTGTGCAAGATGAATGTTGGCCGAGTGTATCACTCATGTTATGAGGACGTCGATAT	3178
Dp	4403	GTGTGCAAGATGAATGTTGGCCGAGTGTATCACTCATGTTATGAGGACGTCGATAT	4462
OY	3179	TCTCTTACTGTCATGCCCATCCGTAAGATGCTTTTCTGTAGCTGTGATGACTCAACCAAG	3238
Dp	4463	TCTCTTACTGTCATGCCCATCCGTAAGATGCTTTTCTGTAGCTGTGATGACTCAACCAAG	4522
OY	3239	TCAATTCGAGAAATGCTATGAGGGCGACCGAGTTGCTCTTGCCCGGCTCAATACGGGAT	3298
Dp	4523	TCAATTCGAGAAATGCTATGAGGGCGACCGAGTTGCTCTTGCCCGGCTCAATACGGGAT	4582
OY	3299	AATACCGGCGCACATAGCAGAACTTTAAAGTGCATCATTTGAAAACGTTCTTCGGGG	3358
Dp	4583	AATACCGGCGCACATAGCAGAACTTTAAAGTGCATCATTTGAAAACGTTCTTCAGAGG	4642
OY	3359	CGAAAACTCTCAAGGATCTTACCGCTGTGAGATCAGTGTGATGATACCACTCGTGTCA	3418
Dp	4643	CGAAAACTCTCAAGGATCTTACCACTGTGAGATCAGTGTGATGATACCACTCGTGTCA	4702
OY	3419	CCCAACTGATCTTGACATCTTTTACCTTCAACAGCGTTTCTGGGTGAGCAAAAAACAGA	3478
Dp	4703	CCCAACTGATCTTGACATCTTTTACCAACAGCGTTTCTGGGTGAGCAAAAAACAGA	4762
OY	3479	AGGCAGAAATGCGCGCAAAAAAGGGAATAAAGGCGACACGGAAATGTGAAATCTCACTACTC	3538
Dp	4763	AGGCAGAAATGCGCGCAAAAAAGGGAATAAAGGCGACACGGAAATGTGAAATCTCACTACTC	4822
OY	3539	TTCCCTTTTCAATATTTATTTGAAGCAATTATCAGGGGTTATGTCTCATGAGGAGATPACATA	3598
Dp	4823	TTCCCTTTTCAATATTTATTTGAAGCAATTATCAGGGGTTATGTCTCATGAGGAGATPACATA	4882
OY	3599	TTTGAATGTATTTAGAAAAATPAACAAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTG	3658
Dp	4883	TTTGAATGTATTTAGAAAAATPAACAAATTTCCGCGCACATTTCCCGCAAAAGTG	4942
OY	3659	CCAC 3662	
Dp	4943	CCAC 4946	

DEFINITION	Accession
Artificial sequences, plasmid vector pGreen-2.	Y09374
Y09374.1 GI:1684627	
blatEM gene; lacZ-phoc fusion gene; multiple cloning site; T3 promoter; T7 promoter.	
synthetic construct	
synthetic construct	
artificial sequences.	
1	
Thaller,M.C., Berlucchi,F., Schippa,S., Selan,L. and Rossolini,G.M.	
Bacterial acid phosphatase gene fusions useful as targets for cloning-dependent insertional inactivation	
Biotechnol. Prog. 14 (2), 241-247 (1998)	
98215688	
JOURNAL	
MEETING	
PUBMED	
9548775	
2 (bases 1 to 3633)	
Rossolini,G.M.	
Direct Submission	
Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia	
Molecular-sez., Microbiologia,, Univ. di Siena, via Laterina N.8,	
1- 53100 Siena, ITALY	
Related sequences: X64444 and X52328.	
Location/Qualifiers	
1. .3633	
/organism="synthetic construct"	
/mol_type="other DNA"	
/db_xref="taxon:32630"	
/plasmid="pGreen-2"	
129. .989	
/gene="blatEM"	
129. .989	
/gene="blatEM"	
/codon_start=1	
/transl_table=11	
/protein_id="CAA70549.1"	
/db_xref="GI:1684628"	
/translation="MSIQRRVALLIPFAFCIPVPAHPETLVKQDAEQDARGCY	
IEIDNSGKLTESFREERPPMSTFKVLGCVLSRIDGQQLRRHHYSONDVE	
YSVPTEKHITDGMIVELCSAALTMSDNTAANLLTTTIGPKELTAFLNNMGHVRLL	
DRNEPELNEAIIPNDRDPTVPVMAATTLRLKLTGLGELLASRQOLLIDMEADVAGPL	
LRSLALPGWFIADKSGGERGSRGIIAALGPDKSPRIVITYTTSQATWIDENRQIA	
ETASLTKHW"	
2147. .2162	
/note="PUC M13/Reverse Sequencing primer recognition site"	
2154. .3104	
/gene="lacZ-phoc"	
2154. .3104	
/gene="lacZ-phoc"	
/codon_start=1	
/transl_table=11	
/protein_id="CAA70550.1"	
/db_xref="GI:1684629"	
/translation="MTMTSPASQLTLTKNKSWSSTAVAALELVDPGCRNISTLS	
IPSTSRGQVPNPSPYSESYTARSLAVVLQRDRENPGVTHGALIPANDTATTKRDLYLL	
KNQQAIDSLKLLRPPEVGSIQFLNDQAMYEKRMRLNTERGKQAQADADLAAGVAT	
AFSGARQYPTTEKSDPELYKLLTNMTEDAGDLATBSAKHMYIRPAPYGETCTCK	
DDKKLTNGSVSPSGTISIGWATLVLAENVANQADAILRGVQLGGSRYVLCGYHMOSD	
VDAAIRYSAVAVTLTSDPAFQAQLAKAQEFQKQKDL"	
2179. .2198	
/gene="lacZ-phoc"	
/note="T3 promoter"	
2214. .2316	
/gene="lacZ-phoc"	
/note="multiple cloning site"	
2325. .2344	
/gene="lacZ-phoc"	
/note="T7 promoter"	
2354. .2370	
/gene="lacZ-phoc"	
/note="M13-20 sequencing primer recognition site"	
2374. .2390	
/gene="lacZ-phoc"	
/note="M13-40 sequencing primer recognition site"	

ORIGIN

Query Match 69.7%; Score 2550.6; DB 12; Length 3633;
 Best Local Similarity 82.3%; Pired. No. 0;
 Matches 3020; Conservative 0; Mismatches 604; Indels 47; Gaps 6;

QY 1 CTAAATTTGTAAGCGTTAATATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAAGCTC 60
 DB 3633 CTAAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAAGCTC 60
 QY 61 ATTTTAAACCAATAGCGCGGAATCGGCAAAATCCCTATATATCAAAAGATAGACCGA 120
 DB 3573 ATTTTAAACCAATAGCGCGGAATCGGCAAAATCCCTATATATCAAAAGATAGACCGA 120
 QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGACAAGATCCACTATTTAAAGACGTGACTC 180
 DB 3513 GATAGGGTTGAGTGTGTTCCAGTTTGGACAAGATCCACTATTTAAAGACGTGACTC 180
 QY 181 CAACGTCAAGGGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTAGCAACCATCAC 240
 DB 3453 CAACGTCAAGGGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTAGCAACCATCAC 240
 QY 241 CTAAATCAAGTTTGTGGGGTGAAGTCCGTTAAAGCACTAATCGGAACCTTAAAGGAG 300
 DB 3393 CTAAATCAAGTTTGTGGGGTGAAGTCCGTTAAAGCACTAATCGGAACCTTAAAGGAG 300
 QY 301 CCCCCGATTAGAGCTTGAACGCGGGAAGCCGCGCAAGTGGCGAAGAAAGGAGGAGAA 360
 DB 3333 CCCCCGATTAGAGCTTGAACGCGGGAAGCCGCGCAAGTGGCGAAGAAAGGAGGAGAA 360
 QY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGCGTCAAGCTGCGGTAACAC 420
 DB 3273 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGCGTCAAGCTGCGGTAACAC 420
 QY 421 CACACCCGCGCGCTTAATGCGCGCTTACAGGCGCGTCCCATTCGCAATTCAGGCTGCG 480
 DB 3213 CACACCCGCGCGCTTAATGCGCGCTTACAGGCGCGTCCCATTCGCAATTCAGGCTGCG 480
 QY 481 CAACGTGGAAGAGCGCATCGGTGCGGGCCCTCTCCGTTATTAAGCCAGCTGGCGAAG 540
 DB 3153 CAACGTGGAAGAGCGCATCGGTGCGGGCCCTCTCCGTTATTAAGCCAGCTGGCGAAG 540
 QY 541 GGGATGTGCTGAAGCGCATTAAGTGGGTAAACCCAGGAGTTTCCAGTCAAGAGCTTG 600
 DB 3093 TCTGTGATTTTGTCAAAATTCGTTTGGCTTTCGTAATCGGCTGAAATGCGGAGT 600
 QY 601 TAAACGACGCGCAGTGAAGCGCGCTCGTTCAATTCAGTTTGAACCCGTGAGAGACG 660
 DB 3033 CGGAATGTAAATGTGCGGACAGCGGCTGAACCGCAATCCGCGGCGATCACAATCACT 660
 QY 661 GCAAGCTCGCGGTGCAAAATGTGTTTACAGCGTGAAGAGAGTGAAGATCTGACAC 720
 DB 2973 GCGAGTGAATGCGCGCAATCAACCC-----GCGTGTGCGAGCTGATTAACCC 720
 QY 721 GCTCAGAAACGCGAGCTGAATTAACCTTAAGAAAGTATCATTTGTGACGTACGTAA 780
 DB 2925 GTTCAGAAATGCAATCTGATTTGCGGGTTCACTTCGCGCACTAATGCGGTGCGC 2865
 QY 781 AGATAATCAGCGTAAATTTGACGATGGGATCTGAATACAGCTCACTATAGGCGAAT 840
 DB 2865 AGCGATAGACGTATACCTGAACGGGTAGATCCGTGTGTGAGAGTTTTCGTGATCTT 2806
 QY 841 TGGGTACCGGGCCCCCTCGAGGTGAACGATGATGATGATGATGATGATGATGATGAT 900
 DB 2805 TGGGTATTAAGGTTTCTGTCCGTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 900
 QY 901 AGCCCGGGGAGATCCACTAATTTTAAGCGCGCGCAACCGCGAGAGCTCAAGTTTGT 960
 DB 2745 TGGCGGAGCGGTTGCGAGGTGACCGGCGATCTCAATCATATGTGTGACAGATTAAACA 2686
 QY 961 TCCCTTTAGTGAAGGTAAATTAAGATCCATGCGTCATTTTACGACAGATATCTTTCTAG 1020
 DB 2685 GCTCCGAGAGTCTTTTTCGTTATCGATGCGGAATG--CCCTGAAATAATCGGTTGC 2628

QY 1021 GGTAAATAGCTGATCAGATATATCGTGGGCTTTTTCGCGCTCAAGCATCGCC 1080
 DB 2627 CACCCCCCTCGGCGAGGTGACAGATCTGCTGTGCTGTTTTCGCGCTCGGATTTGCG 2568
 QY 1081 CAAGCTGCGCTATCTGGGATCGGGAGAGAAACCCGTGCTTTTCCGCGAGTTG 1140
 DB 2567 CAGCATAGCGGCTTTTCATACATTCGCTGATATTTAAACCTGAATATGCGCACTTC 2508
 QY 1141 AAGCG- CATGAAAGAGTTTCCGAGATGACTGTGCTGATTAAGGAGGAA 1199
 DB 2507 CCGGCGGTGCGGTGAACGTTTACGCTGCTGATGCTGTTTATTTTCAATATATA 2448
 QY 1200 CGACGTTTACATGATGATTCGGGAAGGTGAGCAATGACGCTTTAAAGGTAAC 1259
 DB 2447 ATCCGCTGTGTGTGATCGTTCGCGGATGCTCTCATGGGTAAACGCGAGGTT 2388
 QY 1260 TTGCTTCAAGCCACCTGGGATACAGTTTGTGCGGCTTTTCCGGAACAGTTCCGAG 1319
 DB 2387 TTCCAGTCAAGACGTTGTTAAACGACCGCAGTGAAGCGCGGTAAATACACTACATA 2328
 QY 1320 GTACGCGCGAAGCGCATGACAAACCGAACAATACCGGCAAGCGGAACTGCGTCC 1379
 DB 2327 G-----GCGAATTGGGTACCGGCGCCCCCTCGAGGTGACGCTATCATAAGCT 2277
 QY 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGAGAGAGGGTAT 1439
 DB 2276 TGAATTCAAATTCCTGACGCCCGGGGATCACAATTGTTAAGCGGCCGCCACCGG 2217
 QY 1440 CTTGCTGATGATCGGCAAAATGACATGATATCCCGTGAATTAACCGGCGCGCT 1499
 DB 2216 GAGGCT-----CAGCTTTTGTTCCTTTAGTGAAGGTTAATTTGGCGCT 2172
 QY 1500 TGGCGTATCATGCTGCTTACCTGTTCTGTGTGAATTTGTTATCGCTCAATTTGAC 1559
 DB 2217 TGGGTATCATGCTGCTTACCTGTTCTGTGTGAATTTGTTATCGCTCAATTTGAC 2112
 QY 1560 ACAACATACAGCGGAGAGCATTAAGTGAAGCGTGGGGTCCCTAATGATGAGTAA 1619
 DB 2111 ACAACATACAGCGGAGAGCATTAAGTGAAGCGTGGGGTCCCTAATGATGAGTAA 2052
 QY 1620 TCAATTAATTTGCTGCTGCTCACTGCGCGCTTTTCACTGCGGAAACCTGTGCGGAG 1678
 DB 2051 TCAATTAATTTGCTGCTGCTCACTGCGCGCTTTTCACTGCGGAAACCTGTGCGGAG 1678
 QY 1679 -----CTGATTAATGATGCGCAACGCGCGGAGAGAGCGGTTTGGGTATG 1932
 DB 1991 TGCATGCTGATTAATGATGCGCAACGCGCGGAGAGAGCGGTTTGGGTATG 1932
 QY 1732 CTCTCCGCTTCTGCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1791
 DB 1931 CTCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1872
 QY 1792 ATCAGCTCACTCAAGGCGGTAAATACGTTATCCAGAGATTCAGGGATTAACGAGAA 1851
 DB 1871 ATCAGCTCACTCAAGGCGGTAAATACGTTATCCAGAGATTCAGGGATTAACGAGAA 1812
 QY 1852 GAACATGAGCAAAAGCGCAAGAAAGCGCAAGAAACCGTAAAGAGCGCGGTTGCTG 1911
 DB 1811 GAACATGAGCAAAAGCGCAAGAAAGCGCAAGAAACCGTAAAGAGCGCGGTTGCTG 1752
 QY 1912 GTTTTTCATAGGCTCGGCCCCCTGAGCGCATCAAAATTCAGCTCAAGTCAAG 1971
 DB 1751 GTTTTTCATAGGCTCGGCCCCCTGAGCGCATCAAAATTCAGCTCAAGTCAAG 1692
 QY 1972 GTGCGAAACCGCAAGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2031
 DB 1691 GTGCGAAACCGCAAGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1632
 QY 2032 GCGCTCTCTGTTCCAGCTGCGCTTAACGGGATTAACGGGATTAACGGGATTAACGG 2091
 DB 1631 GCGCTCTCTGTTCCAGCTGCGCTTAACGGGATTAACGGGATTAACGGGATTAACGG 1572

OY	2092	TAAAGCGTGGCGCTTTCTCATAGCTCAACGCTGTAGGTATCTCAAGTTGGGTGTAGTGGTTCG	2157
Db	1571	AAGCGTGGCGCTTTCTCATAGCTCAACGCTGTAGGTATCTCAAGTTGGGTGTAGTGGTTCG	1511
OY	2152	CTCCAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTACAGCCGACCGGTGGCGCTTATCCGG	2211
Db	1511	CTCCAAAGCTGGGCTGTGTGCAAGAAACCCCGCTTACAGCCGACCGGTGGCGCTTATCCGG	1451
OY	2212	TAACTATCGTCTTGAAGTCCAAACCCGGTAAAGACAGACTTATCCGCACTGGCAGCAGCAC	2277
Db	1451	TAACTATCGTCTTGAAGTCCAAACCCGGTAAAGACAGACTTATCCGCACTGGCAGCAGCAC	1397
OY	2272	TGGTAAACAGATTACAGAGCGAGGTATGTAGCGCGTGTCTACAGACTTCTGAAGTGTG	2331
Db	1391	TGGTAAACAGATTACAGAGCGAGGTATGTAGCGCGTGTCTACAGACTTCTGAAGTGTG	1332
OY	2332	GCCTAACACGCGCTACATAGAAAGGACAGATATTGGTATCTGGCGTCTGGCTGAAGCCAGT	2397
Db	1331	GCCTAACACGCGCTACATAGAAAGGACAGATATTGGTATCTGGCGTCTGGCTGAAGCCAGT	1277
OY	2392	TACCTTCGGAAAAAGAGTTGGTAGTCTTGGATCCGGCAACAAACACCGCTGTAGCGG	2457
Db	1271	TACCTTCGGAAAAAGAGTTGGTAGTCTTGGATCCGGCAACAAACACCGCTGTAGCGG	1212
OY	2452	TGGTTTTTTTTTGTTCGAAAGCAGCAGATTACGGCGCAAAAAAAGATCTCAAGAGATCC	2511
Db	1211	TGGTTTTTTTTTGTTCGAAAGCAGCAGATTACGGCGCAAAAAAAGATCTCAAGAGATCC	1152
OY	2512	TTTGATCTTTTCTACGGGGTCTGAGGCTCAGTGGAAACGAAACCTACGTTAAGGATTTT	2571
Db	1151	TTTGATCTTTTCTACGGGGTCTGAGGCTCAGTGGAAACGAAACCTACGTTAAGGATTTT	1092
OY	2572	GCTCAGAGATTATCAAAAAAGAGATCTTACACTAGATCTTTTAAATTAAAAATGAAGTTT	2631
Db	1091	GCTCAGAGATTATCAAAAAAGAGATCTTACACTAGATCTTTTAAATTAAAAATGAAGTTT	1032
OY	2632	TAAATCAATCTTAAGTATATATAGTAAACTTGGTCTACAGATTACCAATGCTTATACAG	2691
Db	1031	TAAATCAATCTTAAGTATATATAGTAAACTTGGTCTACAGATTACCAATGCTTATACAG	972
OY	2692	TGAGGACACTTATCAGGAGTCTGTCTAATTCGTTGCAATCCATAGTTGGCTACATCCCGGT	2751
Db	971	TGAGGACACTTATCAGGAGTCTGTCTAATTCGTTGCAATCCATAGTTGGCTACATCCCGGT	912
OY	2752	CGTGTAGATTAATCTACAGATACGGGAGGCTTACCATCTGGACCCAGTGTGCAATGATAC	2811
Db	911	CGTGTAGATTAATCTACAGATACGGGAGGCTTACCATCTGGACCCAGTGTGCAATGATAC	852
OY	2812	CGAGAGCCACGCTCAACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGC	2871
Db	851	CGAGAGCCACGCTCAACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGC	792
OY	2872	CGAGGCGAAGTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGGCG	2931
Db	791	CGAGGCGAAGTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGGCG	732
OY	2932	GGAAGCTAGATTAAGTATGTGCGCAGTTAATAGTTTGGCAACGTTGTGCAATGCTAC	2991
Db	731	GGAAGCTAGATTAAGTATGTGCGCAGTTAATAGTTTGGCAACGTTGTGCAATGCTAC	672
OY	2992	AGGCATCGTGTGTCAACGCTCGTGTGTGTATGGCTTCAATTCAGCTCCGGTCCCAAG	3051
Db	671	AGGCATCGTGTGTGTCAACGCTCGTGTGTGTATGGCTTCAATTCAGCTCCGGTCCCAAG	612
OY	3052	ATCAAGGGAATTAATCAATGATCCCAATGTTGTGCAAAAAAAGCGTTAGCTCTTCGGTCC	3111
Db	611	ATCAAGGGAATTAATCAATGATCCCAATGTTGTGCAAAAAAAGCGTTAGCTCTTCGGTCC	552
OY	3112	TCCGATCGTGTGCAAGATAGTTGGCGCAGTGTATCACTCAATGGTTATGGCAGCACT	3171
Db	551	TCCGATCGTGTGCAAGATAGTTGGCGCAGTGTATCACTCAATGGTTATGGCAGCACT	492
OY	3172	GCAATATCTCTACTGTCAATGCAATCCGTAAGATGCTTTTCTGTGACTGTGTAGATAC	3231

Db	Query Match	Best Local Similarity	Matches 3052;	Conservative	Score 2517.8;	DB 6;	Length 3918;	Indels 142;	Gaps 7;
Db	491	GCATATATCTCTTACTGTGATGCGATCCGTAAGTGCCTTTCTGTGA	CGTGGAGTACTC	432					
Qy	3232	AACCAAGTCATTTCTGAGATAGTGTATGCGGGACCGAGTTGCTTTG	CCCGCGCTCAAT	329					
Db	431	AACCAAGTCATTTCTGAGATAGTGTATGCGGGACCGAGTTGCTTTG	CCCGCGCTCAAT	372					
Qy	3292	ACGGGATATATACCCGCGCCACATAGCAGAACTTTTAAAGTGTCA	TATTGGAAAACTTTC	3351					
Db	371	ACGGGATATATACCCGCGCCACATAGCAGAACTTTTAAAGTGTCA	TATTGGAAAACTTTC	312					
Qy	3352	TTTCGGGGGCGAAAACTCTCAAGGATTTTACCGTGTGATCCAGTT	CGATGTAAACCAAC	3411					
Db	311	TTTCGGGGGCGAAAACTCTCAAGGATTTTACCGTGTGATCCAGTT	CGATGTAAACCAAC	252					
Qy	3412	TGCTGACCCCACTGATCTTCAGCATCTTTTACTTTTCAACAAGCT	TTCGGGTGACAAA	3471					
Db	251	TGCTGACCCCACTGATCTTCAGCATCTTTTACTTTTCAACAAGCT	TTCGGGTGACAAA	192					
Qy	3472	AACAGAAAGCGAAAAATGCGCGAAAAAAGGAATTAAGCGCGAC	ACAGAAATGTTGAATCT	3531					
Db	191	AACAGAAAGCGAAAAATGCGCGCGAAAAAAGGAATTAAGCGCG	ACAGAAATGTTGAATCT	132					
Qy	3532	CATACCTCTTCTTTTCAATATATATGAGCAATTAATCAGGTTAT	TGTCTCATGACGCG	3591					
Db	131	CATACCTCTTCTTTTCAATATATATGAGCAATTAATCAGGTTAT	TGTCTCATGACGCG	72					
Qy	3592	ATGACATTTTGAATGATTTTGAAGAAAAAATTAACAAATAGGG	GTTCGCGCAATTTCCCG	3651					
Db	71	ATGACATTTTGAATGATTTTGAAGAAAAAATTAACAAATAGGG	GTTCGCGCAATTTCCCG	12					
Qy	3652	AAAAAGTGCAC	3662						
Db	11	AAAAAGTGCAC	1						
RESULT 13									
LOCUS	CQ768840/c								
DEFINITION	Sequence 59 from Patent WO2004007550.								
ACCESSION	CQ768840								
VERSION	CQ768840.1								
KEYWORDS									
SOURCE									
ORGANISM									
REFERENCE									
AUTHORS	1 Unger, C.M., Zehetmayer, C., Torella, C., Niewoehner, J., Ahrens, B. and Beste, G.								
TITLE	Mcam inhibitors								
JOURNAL	Patent: WO 2004007550-A 59 22-JUN-2004;								
FEATURES									
source									
ORIGIN									
Query Match	68.8%;	Score 2517.8;	DB 6;	Length 3918;					
Best Local Similarity	81.8%;	Pred. No. 0;							
Matches 3052;	Conservative	0;	Mismatches 537;	Indels 142;	Gaps 7;				
Qy	3	AAATTTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGT	TTAAATCAGCTCAT	62					
Db	3732	AAATTTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGT	TTAAATCAGCTCAT	3673					
Qy	63	TTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTTAAATCAAA	GAATAGACGAG	122					
Db	3672	TTTTTAACCAATAGCGCGAAATCGGCAAAATCCCTTTAAATCAAA	GAATAGACGAG	3613					
Qy	123	TAGGGTTGAGTGTGTTTCCAGTTTGGAAACAAGATCCACTTAA	AGAAAGGTGATCCCA	182					

Db 3612 TAGGGTTAGTATTTGTTCCAGTTTGGAAACAAGATCCACTATTAAGAAAGTGAGACTCGA 3553
Qy 183 AGCTGAAGGCGGAAACCCGTCATCAAGGCGATGCGCCACTACGGAACCATCACCT 242
Db 3552 AGGTAAAGGACGAAAAACCGTCATCAAGGCGATGCGCCACTACGGAACCATCACCT 3493
Qy 243 AATCAAGTTTTTTGGGGGTCAGAGTGCCTTAAGCACTAAATCCGAACCTTAAAGGAGCC 302
Db 3492 AATCAAGTTTTTTGGGGGTCAGAGTGCCTTAAGCACTAAATCCGAACCTTAAAGGAGCC 3433
Qy 303 CCCGATTAGACTTGAACGAGGAAACCGGCGAAACGTGCGAGAAAGGAAAGAAAG 362
Db 3432 CCCGATTAGACTTGAACGAGGAAACCGGCGAAACGTGCGAGAAAGGAAAGAAAG 3373
Qy 363 CGAAGAGCGGCGGCTAAGGCGCTGCGAAAGTGAAGCGTACGCTGCGCGTAACCA 422
Db 3372 CGAAGAGCGGCGGCTAAGGCGCTGCGAAAGTGAAGCGTACGCTGCGCGTAACCA 3313
Qy 423 CACCCGCGGCTTAATGCGCGCTACAGGCGCGT----- 458
Db 3312 CACCCGCGGCTTAATGCGCGCTACAGGCGCGTACTATAGCTTGAAGTACG 3253
Qy 459 ----- 459
Db 3252 GTGTGAATATCCGCAAGATGCGTAAGAAATAACGATCAGGCGCAATTGCGCAAT 471
Qy 472 CAGGCTGCGCACTGTTGGGAAGGCGATCGGTGCGGCGCTTTCGCTAATACGCACT 531
Db 3192 CAGGCTGCGCACTGTTGGGAAGGCGATCGGTGCGGCGCTTTCGCTAATACGCACT 3133
Qy 532 GGGAAAGGGGATGCTGCGCAAGCGCTAATAGTTGGTAAACGCAAGGTTTTCCAGTC 591
Db 3132 GGGAAAGGGGATGCTGCGCAAGCGCTAATAGTTGGTAAACGCAAGGTTTTCCAGTC 3073
Qy 592 ACAGCTGTATAAACGACGCGCAAGTGAAGCGCGCGCTGCTAATACGTTTTGAACCGT 651
Db 3072 ACAGCTGTATAAACGACGCGCAAGTGAAGCGCGCGCTGCTAATACGTTTTGAACCGT 3013
Qy 652 GGAGAGCGGCGAGTCCGCGTGAATGTGTTTTACAGCGTGAAGAGATGAAGAT 711
Db 3012 -----TAGAAACGTGAAATAATACATTAAGTGGCAACATTAAGAA 2964
Qy 712 GCTGACACGCTGAGAACACGAGCTAGATTAACCTAGAAAGATATATGTCAT 771
Db 2963 CGAAAGACACGAGATTAAGTTATTTGTCAATCAATGAATAATTCATATGTTT 2904
Qy 772 GTACGTTAAAGATATATGCTAATAATTAAGCATGCGATCTGTATATGACTCAT 831
Db 2903 ACCAGCGCCAAAGCATTAAGGCGACATTCACCGACTGAGCAAGAAAGTAAATATG 2844
Qy 832 AGGCGAATTTGGGTACCGGCGCGCGCTCGAGGTGACGATACGATTAAGTTATG 891
Db 2843 ACCGAAATTTATCATTAAGTGAATTTACCGCTACGACCTTGAATTTGGAAT 2784
Qy 892 AATTCCGCGACCGCGGAGATCACTAGTTCTAAGCGCGCGCGCAACCGGCGTGAAGTC 951
Db 2783 AGAGCCAGCA-----AATTCACCAATGACACCATTTACATTTAGCAAGCGGAAACG 2732
Qy 952 AGCTTTGTTCCCTTATAGTGGGTTAATTAAGTCCCATGCGCTCAATTTTACGAC 1011
Db 2731 TCACCAATGAACCATGATAGACAGCCGTATATCAATGACAGCAAGATCA---AGTTG 2675
Qy 1012 TCTTTTAAAGGTTATCTAGTGCATCAAGATCATATGCGGCGCTTTTTTTCGAGCTCA 1071
Db 2674 CTTTAGCTCAAGACTGACGCGTTTTCATCGCATTTTTCGTCATAGCCCCCTTATTA 2615
Qy 1072 GTCATGCGCCAAGCTGCGCTATCTGGGATCTGGGAGAAAGAACCGCGCTTTTCCC 1131
Db 2614 GGGTTGCGCATTTTTCATATCAAAATCACCGGAACCGAGCGCGCAACGGAACGCGCA 2555
Qy 1132 GGAAGTTGAAGCGCGATGAAAGATTTGCGGAGATGAAGTGTGCTGATTTGACGTTG 1191

Db 2554 CCTCAGAGCGCCACCTCTAGAACCGCACCGCTCAGAGCGCCACCTCTAGAGCGCCA 2495
Qy 1192 AGCGAAACCGAGCTTTTACATGATGATTTGGGAAGGTGTGGCATGACGCTTTAAACG 1251
Db 2494 CCABAACACAC-----CAGAGCGCGCGCAGCTTATCGGCGACGCGGTCCAGGAT 2440
Qy 1252 GTGAATGTTGTTTCAAGGCGACCTGGGATACAGTTGCTGCGGCGTTTTCCGGAACAGT 1311
Db 2439 CCGGATACGCGCACCGCGCACCTGCGCGCGCTGTTGATGATGACCTTCGCGCACGCGC 2380
Qy 1312 TCCGATGCTCAGCCCGGAAGCGCATGACAAACCGGAACATTAACGCGACAGCCGAACT 1371
Db 2379 AGAGCACCTCGCCTGAAACCGCTTCACCTGTGCA----- 2344
Qy 1372 GCCGTGCGGTGTGAGATTAATGAACAGCGGTGCGCGCTGGGATTAATACGTACGAG 1431
Db 2343 -----CTGACCTTGGGCGATGCGCGGTGGCGCGCATTAAGAAACAACTAAGAA 2291
Qy 1432 ACGGATATCTGGCTGGAATGCGCGAAATGGAATGGAATCCCGTGAAGTTACCGGCG 1491
Db 2290 TTGCGAATATATTTTTCAGTTGAAATCTCAAAAAA-----GCT 2244
Qy 1492 GGGCGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
Db 2243 CCAAGCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2184
Qy 1552 AATTCCACACAAATATGACGCGGAAAGATTAAGTTAAAGCTGGGCTCTAATGAT 1611
Db 2183 AATTCCACACAAATATGACGCGGAAAGATTAAGTTAAAGCTGGGCTCTAATGAT 2124
Qy 1612 GAGCTAATCAATTAATTTGCTGCTCACTGCGCGCTTTCAGTGGGAAACCTGTC 1671
Db 2123 GAGCTAATCAATTAATTTGCTGCTCACTGCGCGCTTTCAGTGGGAAACCTGTC 2064
Qy 1672 GTGCGAGTGTATTAAGATTCGCGCAACGCGGGAAGAGCGGTTTTGCTGATTTGGGCG 1731
Db 2063 GTGCGAGTGTATTAAGATTCGCGCAACGCGGGAAGAGCGGTTTTGCTGATTTGGGCG 2004
Qy 1732 CTCTTCGCTTCTGCTGCTCACTGACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1791
Db 2003 CTCTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
Qy 1792 ATGAGCTCACTCAAGCGGTAATACGTTATCAAGATCAGGGAATTAACGAGAA 1851
Db 1943 ATGAGCTCACTCAAGCGGTAATACGTTATCAAGATCAGGGAATTAACGAGAA 1884
Qy 1852 GAAATGTGAGCAAAAGCGAGCAAAAGCGAGAAAGCGTAATAAGCGCGGTTGCTGCG 1911
Db 1883 GAAATGTGAGCAAAAGCGAGCAAAAGCGAGAAAGCGTAATAAGCGCGGTTGCTGCG 1824
Qy 1912 GTTTTTCATAGGCTTCGCGCTGAGAGCATCAAAATATGACGCTCAATCAAG 1971
Db 1823 GTTTTTCATAGGCTTCGCGCTGAGAGCATCAAAATATGACGCTCAATCAAG 1764
Qy 1972 GTGGCGAAACCGGAGAGATTAAGATACAGGCGTTTCCCTGGAAGCTCCCTGT 2031
Db 1763 GTGGCGAAACCGGAGAGATTAAGATACAGGCGTTTCCCTGGAAGCTCCCTGT 1704
Qy 2032 GGGCTCTCTGTTCCGACCTTGGCGTTACCGGATACCTGTCCGCTTTTCTCCCTTGCG 1644
Db 1703 GGGCTCTCTGTTCCGACCTTGGCGTTACCGGATACCTGTCCGCTTTTCTCCCTTGCG 1644
Qy 2092 AAGCTGCGCTTCTCAATAGCTACGCTGATAGATATCAAGTTCCGTTGATGATGATGAT 2151
Db 1643 AAGCTGCGCTTCTCAATAGCTACGCTGATAGATATCAAGTTCCGTTGATGATGATGAT 1584
Qy 2152 CTCCAAAGCTGCTGTGACGAACCCCGTTACAGCGCGGACGCGCTTATCCG 2211
Db 1583 CTCCAAAGCTGCTGTGACGAACCCCGTTACAGCGCGGACGCGCTTATCCG 1524
Qy 2212 TAACTATGCTTGAAGTCAACCGGTAAGACAGCTTATGCGCATGCGAGACGAC 2271
Db 1523 TAACTATGCTTGAAGTCAACCGGTAAGACAGCTTATGCGCATGCGAGACGAC 1464

QY 2272 TGGTAAAGAGATTAGCAGAGGAGGTATGTAGCGCGTCTACAGAGTTCTTGAAGTGTG 2331
 DB 1463 TGGTAAAGAGATTAGCAGAGGAGGTATGTAGCGCGTCTACAGAGTTCTTGAAGTGTG 1404
 QY 2332 GCGTAATACGCGCTACACTAGAGAGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGT 2391
 DB 1403 GCGTAATACGCGCTACACTAGAGAGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGT 1344
 QY 2392 TACCTTCGGAAGAAAGATTGGTAGTCTTGAATCCGGCAACAAACACCGCTGTGAGCGG 2451
 DB 1343 TACCTTCGGAAGAAAGATTGGTAGTCTTGAATCCGGCAACAAACACCGCTGTGAGCGG 1284
 QY 2452 TGGTTTTTTTGGTGAAGCAGCAGATTAACGGCGGAGAAAAAGAAATCTCAAGAAAGCC 2511
 DB 1283 TGGTTTTTTTGGTGAAGCAGCAGATTAACGGCGGAGAAAAAGAAATCTCAAGAAAGCC 1224
 QY 2512 TTTGATCTTTTCTACGCGGCTGACGCTCAGTGAACGAAAACTCAGCTTAAGGATTTT 2571
 DB 1223 TTTGATCTTTTCTACGCGGCTGACGCTCAGTGAACGAAAACTCAGCTTAAGGATTTT 1164
 QY 2572 GGTGATGATTTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATAAGTATTT 2631
 DB 1163 GGTGATGATTTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATAAGTATTT 1104
 QY 2632 TAAATCAATCTAAAGTATATATGAGTAACTTGGCTGACAGTTACCAATGCTTAATAG 2691
 DB 1103 TAAATCAATCTAAAGTATATATGAGTAACTTGGCTGACAGTTACCAATGCTTAATAG 1044
 QY 2692 TGAGGCACTTATCTCAGCGATCTGTCTATTTGTTTCATTCATAGTGGCTGACTCCCGT 2751
 DB 1043 TGAGGCACTTATCTCAGCGATCTGTCTATTTGTTTCATTCATAGTGGCTGACTCCCGT 984
 QY 2752 CGTGTAGTAACTACGATACGGAAGGAGGCTTACATCTGGCCCCAGTGTGCAATGATACC 2811
 DB 983 CGTGTAGTAACTACGATACGGAAGGAGGCTTACATCTGGCCCCAGTGTGCAATGATACC 924
 QY 2812 GCGAGACCCAGCTCACCGGCTCCAGATTTATACAGCAATAAACGACGCGGAGAGGCG 2871
 DB 923 GCGAGACCCAGCTCACCGGCTCCAGATTTATACAGCAATAAACGACGCGGAGAGGCG 864
 QY 2872 CGAGCGGAGAGTGGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCG 2931
 DB 863 CGAGCGGAGAGTGGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCCG 804
 QY 2932 GGAAGCTAGAGTAAAGTATGCGCAGTTAATAGTTTGGCAACGTTGTCATTGCTAC 2991
 DB 803 GGAAGCTAGAGTAAAGTATGCGCAGTTAATAGTTTGGCAACGTTGTCATTGCTAC 744
 QY 2992 AGGATCGTGTGTACAGCTGCTGTTTGGTATAGGCTTACAGCTCCGCTTCCCAAG 3051
 DB 743 AGGATCGTGTGTGTACAGCTGCTGTTTGGTATAGGCTTACAGCTCCGCTTCCCAAG 684
 QY 3052 ATCAAGGCGAGTTCATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCCGCTC 3111
 DB 683 ATCAAGGCGAGTTCATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCCGCTC 624
 QY 3112 TCCGATCGTGTGTCAAGAGTAAAGTTGGCGCAGTGTATCACTCAATGTTATGAGCAGCACT 3171
 DB 623 TCCGATCGTGTGTCAAGAGTAAAGTTGGCGCAGTGTATCACTCAATGTTATGAGCAGCACT 564
 QY 3172 GCATTAATTTCTTATCTGTACAGCAATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTC 3231
 DB 563 GCATTAATTTCTTATCTGTGTACAGCAATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTC 504
 QY 3232 AACCAAGTCAATTCAGAAATAGTATGCGGCGAGCCGAGTGTCTTGGCCCGGCGTCAAT 3291
 DB 503 AACCAAGTCAATTCAGAAATAGTATGCGGCGAGCCGAGTGTCTTGGCCCGGCGTCAAT 444
 QY 3292 ACGGAGTAAATCCGCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGAGAAACGTTCC 3351
 DB 443 ACGGAGTAAATCCGCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGAGAAACGTTCC 384

QY 3352 TTGGGGGGAAGAACTCTCAAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCAC 3411
 DB 383 TTGGGGGGAAGAACTCTCAAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCAC 324
 QY 3412 TCGTGAACCCAACTGATCTTCAGCATCTTTTACTTTTACACGACGCTTCTGGGTAGCAGA 3471
 DB 323 TCGTGAACCCAACTGATCTTCAGCATCTTTTACTTTTACACGACGCTTCTGGGTAGCAGA 264
 QY 3472 AACGGAAGGCAAAATCCGCGCAAAAAAGGAAATAGGCGCGACACGAAATGTTGAATCT 3531
 DB 263 AACGGAAGGCAAAATCCGCGCAAAAAAGGAAATAGGCGCGACACGAAATGTTGAATCT 204
 QY 3532 CATACCTCTCTTTTCAATTAATTTAGAGCATTTACAGGCTATTTGTCATGAGCGG 3591
 DB 203 CATACCTCTCTTTTCAATTAATTTAGAGCATTTACAGGCTATTTGTCATGAGCGG 144
 QY 3592 ATACATATTTGAATGTATTTAGAAAAATTAACAATAGGAGTTCGCGCACATTTCCCG 3651
 DB 143 ATACATATTTGAATGTATTTAGAAAAATTAACAATAGGAGTTCGCGCACATTTCCCG 84
 QY 3652 AAAAGTCCAC 3662
 DB 83 AAAAGTCCAC 73
 RESULT 14
 CQ802124/c 3918 bp DNA linear PAT 05-MAY-2004
 LOCUS CQ802124
 DEFINITION Sequence 109 from Patent WO200400717.
 ACCESSION CQ802124
 VERSION CQ802124.1 GI:47058654
 KEYWORDS
 SOURCE
 ORGANISM
 AUTHORS Unger C.M., Beese G., Jensen K.H., Zehetmeier C., Knauer R.,
 Guelbenzu B.L., Torella C. and Ilag L.U.
 TITLE Proteomic screen to identify disease-related biological
 JOURNAL Molecules and inhibitors thereto
 PATENT Patent: WO 200400717-A 109 22-JAN-2004;
 FEATURES
 source
 Location/Qualifiers
 1..3918
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Description of Artificial Sequence: vector pXp10"
 ORIGIN
 Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
 Best Local Similarity 81.8%; Pred. No. 0;
 Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;
 QY 3 AAATGTAAACGCTTAATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTCAT 62
 DB 3732 AAATGTAAACGCTTAATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTCAT 3673
 QY 63 TTTTAAACAATAGGCGGCAAAATCCCTTAATCAAAAAGAAATAGCGAGA 122
 DB 3672 TTTTAAACAATAGGCGGCAAAATCCCTTAATCAAAAAGAAATAGCGAGA 3613
 QY 123 TAGGTTGAGTGTGTTTCCAGTTTGAACAAGATGCTCAATTAAGAACTGAGCTCA 182
 DB 3612 TAGGTTGAGTGTGTTTCCAGTTTGAACAAGATGCTCAATTAAGAACTGAGCTCA 3553
 QY 183 ACGTCAAAAGGCGGCAAAATCCGCTTATCAAGGCGATGCGCCATCACTGAACTACCTT 242
 DB 3552 ACGTCAAAAGGCGGCAAAATCCGCTTATCAAGGCGATGCGCCATCACTGAACTACCTT 3493
 QY 243 AATCAAGTTTGTGGGTGAGGTGCGTAAAGCACTAAATCGAAACCTTAAGGAGGCC 302
 DB 3492 AATCAAGTTTGTGGGTGAGGTGCGTAAAGCACTAAATCGAAACCTTAAGGAGGCC 3433

QY 303 CCCGATTAGACTTGAACGGGAAACCGGCGAACGTGCGAGAAAGAAAGAAAG 362
Db 3432 CCCGATTAGACTTGAACGGGAAACCGGCGAACGTGCGAGAAAGAAAGAAAG 3373
QY 363 CCAAAAGAGCGGCGCTAGAGCGCTGCGAAGTGAAGCGTGAACCTTGGCGTAA 422
Db 3372 CCAAAAGAGCGGCGCTAGAGCGCTGCGAAGTGAAGCGTGAACCTTGGCGTAA 3313
QY 423 CACCGCGCGCTTAATCCCGCTACAGGCGCGT----- 458
Db 3312 CACCGCGCGCTTAATCCCGCTACAGGCGCGTCTATGGTCTTTGACGTATGCG 3253
QY 459 -----CCCATTCGCATT 471
Db 3252 GTGTGAATAACCGACAGATCCGTAAAGAAATAACCGCATCAGCGCATTCGCATT 3193
QY 472 CAGGCTGGCACTGTTGGGAAGGCGATGCTGCGGCGCTCTTGCCTATTAGCGAGCT 531
Db 3192 CAGGCTGGCACTGTTGGGAAGGCGATGCTGCGGCGCTCTTGCCTATTAGCGAGCT 3133
QY 532 GCGCAAGGGGAGTGTCTCAAGGGGATTAAAGTGGTAAAGCCAGGGTTTCCAGTTC 591
Db 3132 GCGCAAGGGGAGTGTCTCAAGGGGATTAAAGTGGTAAAGCCAGGGTTTCCAGTTC 3073
QY 592 ACACGCTTTAAACGACGCGCGAGTACGCGCTGCTTCACTTCAAGTTTGAACCGGT 651
Db 3072 ACACGCTTTAAACGACGCGCGAGTATTTTAAGACTCTTATTAACGAGTATGT 3013
QY 652 GAGAGACGGGACAGTCCGCGTGCAGAAATGTTTAAACGCTGATGAGCAGATGAAGT 711
Db 3012 -----TAGCAAAAGTGCAGAAATTAATCAATTAAGGTGGCAATTAAGAA 2964
QY 712 GCTCGACAGCTGCAGAACACGACGCTAGATTAACTAGAAAGTAACTAATTTGAC 771
Db 2963 CGCAAGACACACGCAATTAAGTTATTTTGTTCACAAATCAATTAAGAAATTTATTTG 2904
QY 772 GTACGTTAAAGATTAATCAATGCGTAAATTTGACGATGGGATCTGTAATACGACTCAT 831
Db 2903 ACCAGCGCCAAAGACATTAAGGCGCATTTCAACGACTGAGGCAAAAGTAAATTTG 2844
QY 832 AGGCGAATTTGGATCCGGGCCCCCTCGAGTGCAGCGTATCGAATTAACCTTGATG 891
Db 2843 AGGCAATTTATTCATTAAAGTGAATTAATCAACGCTACCGACTTGAACCAATTTGGGAAT 2784
QY 892 AATTCCTGCAGCCGCGGAGTCCACTAGTCTTAAGCGCGCCGACCGCGGTGAGCTCC 951
Db 2783 AGAGCAGCA-----AAATCACAGTAGACCACTTACATTAGCAAGCCGGAACG 2732
QY 952 AGCTTTGTTCCCTTAAAGTGAAGGTTAATTAGATCCCATGCGTCATTTTACGAGACTA 1011
Db 2731 TCACCAATGAACCATGATAGACGACCGTATCATGAGCAACAAATCA--AGTTTG 2675
QY 1012 TCTTTAGAGGTTAATTAAGTCAATCAAGATATATCGTGGGCTTTTTCGGGCTCA 1071
Db 2674 CTTTACGCTCAACACTGATGCGGTTTCAATCGGCAATTTTCGATACAGCCCTTATTA 2615
QY 1072 GTATGCCCAAGCTGCGCTATCTGGGATCGGGAGGAAGAAACCGGCTTTTCC 1131
Db 2614 GCGTTTGCATTTTTCATTAATCAAAATCAACCGGAACCGGACCGCCACCGCA 2555
QY 1132 GCGAGGTTGAAGGCGATGGAAGATTTGCGAAGATGAATGCTGTGCAATTAAGCTTG 1191
Db 2554 CCTCAGAGCCGCAACCTCAAGACCGCACTTCAGAGCGCAACCTTCAGAGCCGCA 2495
QY 1192 AGGAAAGACGAGTTTACATGATGATTCGGGAAGTGGCCATGACGCTTTTAAAG 1251
Db 2494 CCAAGAACACAC-----CAGAGCGCGCGACGCTATGCGCAACGCGTTCCAGCGGAT 2440
QY 1252 GTGAACGTGTTGTTGAGGCACTGGAATACAGTTTCTGCGGCTTTTCGAGCAAGT 1311
Db 2439 CCGGATACGGCACCGGCGCACTGCGCGCGCTGTTGATGAGTCACTTCCGCAACGCC 2380

QY 1312 TCCGATGTCAGCCCGAAGCATACAGCAACCGGAACATACCGGCGACAGCCGGAAGT 1371
Db 2379 AAGACACACTCCGCTGAACCGGCTTCACTCTGCA----- 2344
QY 1372 GCCGTCCGCTGTGACATTAATGACAGCGGTGCGGCTGGGATATTAAGTACGAGAG 1431
Db 2343 -----CTGACCTGGCCATGAGCGGCTGGGCGCATTAAGAAAGAAACATTAAGAA 2291
QY 1432 ACGGATATCTGCTGATGCGCGAAGATGACATGATACCCGTGATTAACCGGCG 1491
Db 2290 TTGGATATATATTTTTCAGCTGAATATTCAGAAAAA-----GGCT 2244
QY 1492 GGGCGCTTGGCGTATCATGATGATGCTGTTCTGTTGTAATTTGATCCGCTAC 1551
Db 2243 CCAAGGCTTGGCGTATCATGATGATGCTGTTCTGTTGTAATTTGATCCGCTAC 2184
QY 1552 AATTCACACACATACAGACCGGAAGTAAGTGAAGCTGGGCTGCTATAGAT 1611
Db 2183 AATTCACACACATACAGACCGGAAGTAAGTGAAGCTGGGCTGCTATAGAT 2124
QY 1612 GAGCTAATCAATTAATTCGCTGCTCACTGCGCTTTCAGTTCGGGAAACCTGTC 1671
Db 2123 GAGCTAATCAATTAATTCGCTGCTCACTGCGCTTTCAGTTCGGGAAACCTGTC 2064
QY 1672 GTGCCAGCTGATTAATTAATGCGCAACCGCGGGAAGAGCGGTTTGGATTTGGCG 1731
Db 2063 GTGCCAGCTGATTAATTAATGCGCAACCGCGGGAAGAGCGGTTTGGATTTGGCG 2004
QY 1732 CTCTTCGCTTCTGCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1791
Db 2003 CTCTTCGCTTCTGCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
QY 1792 ATCAGCTCACTCAAAAGGCGTAATACGTTATCAAGATTCAGGGATTAACGAGAA 1851
Db 1943 ATCAGCTCACTCAAAAGGCGTAATACGTTATCAAGATTCAGGGATTAACGAGAA 1884
QY 1852 GAACTGTGACAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGCGCGCTTCTGTC 1911
Db 1883 GAACTGTGACAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGCGCGCTTCTGTC 1824
QY 1912 GTTTTTCATTAAGGCTCCGCCCCCTGACGATCAACAAATTCAGCTCAAGTCAAG 1971
Db 1823 GTTTTTCATTAAGGCTCCGCCCCCTGACGATCAACAAATTCAGCTCAAGTCAAG 1764
QY 1972 GTGCGAAACCGCAGACGATTAAGATACGAGGCTTCCCGTGAAGTCCCTGT 2031
Db 1763 GTGCGAAACCGCAGACGATTAAGATACGAGGCTTCCCGTGAAGTCCCTGT 1704
QY 2032 GCGCTCCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCCCTGCGG 2091
Db 1703 GCGCTCCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCCCTGCGG 1644
QY 2092 AAGCGTGGCTTTCATAGCTCAGCGTGAAGTATCAAGTTCGCTGCTGCTGCTG 2151
Db 1643 AAGCGTGGCTTTCATAGCTCAGCGTGAAGTATCAAGTTCGCTGCTGCTGCTG 1584
QY 2152 CTCGAAGCGGCGCTGTGACAAACCCCGTTCAAGCCGACCGGCTGCTTATCCG 2211
Db 1583 CTCGAAGCGGCGCTGTGACAAACCCCGTTCAAGCCGACCGGCTGCTTATCCG 1524
QY 2212 TAACTATGCTTGTAGTCAACCGGTAAGACAGATTAATTCGCTGCTGCTGCTG 2271
Db 1523 TAACTATGCTTGTAGTCAACCGGTAAGACAGATTAATTCGCTGCTGCTGCTG 1464
QY 2272 TGGTAAAGGATTAAGACAGAGGATTAAGGCGGTCTACAGATTTCTTGAAGTGG 2331
Db 1463 TGGTAAAGGATTAAGACAGAGGATTAAGGCGGTCTACAGATTTCTTGAAGTGG 1404
QY 2332 GCTTAATCAAGCTACAGTGAAGACAGATTTGATTTGCTGCTGCTGCTGCTG 2391
Db 1403 GCTTAATCAAGCTACAGTGAAGACAGATTTGATTTGCTGCTGCTGCTGCTG 1344
QY 2392 TACCTTGGAAAAAGATTGATGCTTGTATCCGCAAAACACACGCTGTTAGGG 2451

Db 1343 TACCTCGGAAAAAGATTGGTAGCTTGTATCCGGCAACAAACACCGCTGGTAGCG 1284
 Qy 2452 TGGTTTTTTTGTGGCAAGCAGCAGATTACGGCGAAGAAAAAGAGATCTCAAGAGATCC 2511
 Db 1283 TGGTTTTTTTGTGGCAAGCAGCAGATTACGGCGAAGAAAAAGAGATCTCAAGAGATCC 1224
 Qy 2512 TTGATCTTTTCTACGGGGGCTGACGGCTCAGTGGAGAAAGTCAAGCTTAAGGATTTT 2571
 Db 1223 TTTGATCTTTTCTACGGGGGCTGACGGCTCAGTGGAGAAAGTCAAGCTTAAGGATTTT 1164
 Qy 2572 GGTATGATGATTAATCAAAAAGATCTTACCTAGATCTTTTAAATTAATTAAGATTT 2631
 Db 1163 GGTATGATGATTAATCAAAAAGATCTTACCTAGATCTTTTAAATTAATTAAGATTT 1104
 Qy 2632 TAAATCAATCTAAAGATATATAGATAAAGTGGTGTGACAGTTACCAATGCTTAATCAG 2691
 Db 1103 TAAATCAATCTAAAGATATATAGATAAAGTGGTGTGACAGTTACCAATGCTTAATCAG 1044
 Qy 2692 TGAGGACCTATCTCAGGAGATCTGTATTTCTGATCCATAGTGGCTGACCTCCCGT 2751
 Db 1043 TGAGGACCTATCTCAGGAGATCTGTATTTCTGATCCATAGTGGCTGACCTCCCGT 984
 Qy 2752 CGTGTAGATTAATCAAGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATAC 2811
 Db 983 CGTGTAGATTAATCAAGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATAC 924
 Qy 2812 GCGAGACCCACGCTTCACGGCTCCAGATTTATCAGCAATTAACAGCAGCCGAGAGGGC 2871
 Db 923 GCGAGACCCACGCTTCACGGCTCCAGATTTATCAGCAATTAACAGCAGCCGAGAGGGC 864
 Qy 2872 CGAGCGCAGAGTGGTCTCGCAAGCTTATCGGCGTCCATCCAGTCTATTAATTTGTCG 2931
 Db 863 CGAGCGCAGAGTGGTCTCGCAAGCTTATCGGCGTCCATCCAGTCTATTAATTTGTCG 804
 Qy 2932 GGAAGCTAGAGTATAGTTCGCCAGTAAATAGTTTGGCAACGTTTGGCATTGTCTAC 2991
 Db 803 GGAAGCTAGAGTATAGTTCGCCAGTAAATAGTTTGGCAACGTTTGGCATTGTCTAC 744
 Qy 2992 AGGATCGTGTGTCAAGCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3051
 Db 743 AGGATCGTGTGTGTCAAGCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 684
 Qy 3052 ATCAAGCGAGTATCAAGATCCCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3111
 Db 683 ATCAAGCGAGTATCAAGATCCCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 624
 Qy 3112 TCCGATCGTGTGTCAAGATAGTGGCGCAGTGTATCACTCATGTGTGTGTGTGTGTGTGTGTGTGTGT 3171
 Db 623 TCCGATCGTGTGTCAAGATAGTGGCGCAGTGTATCACTCATGTGTGTGTGTGTGTGTGTGTGTGTGT 564
 Qy 3172 GCATTAATTTCTTACTGTATCCATCCGTAAGATGCTTTTCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGT 3231
 Db 563 GCATTAATTTCTTACTGTATCCATCCGTAAGATGCTTTTCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGT 504
 Qy 3232 AACCAAGTCATCTGAGAAATGTGTATGCGCGACCGAGTGTCTTGGCCGGCGTCAAT 3291
 Db 503 AACCAAGTCATCTGAGAAATGTGTATGCGCGACCGAGTGTCTTGGCCGGCGTCAAT 444
 Qy 3292 ACGGATTAATTCGCGCAGCAGATAGCAGAACTTTAAAGTGTCTCATTTGAGAAAGCTTC 3351
 Db 443 ACGGATTAATTCGCGCAGCAGATAGCAGAACTTTAAAGTGTCTCATTTGAGAAAGCTTC 384
 Qy 3352 TTGGGGGGAAGAACTCTCAGAGATCTTACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3411
 Db 383 TTGGGGGGAAGAACTCTCAGAGATCTTACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 324
 Qy 3412 TCGTGACCCCAAGTCATGATCTTCAAGATCTTTTACCTTACCGAGCTTCTGGGTGAGGAA 3471
 Db 323 TCGTGACCCCAAGTCATGATCTTCAAGATCTTTTACCTTACCGAGCTTCTGGGTGAGGAA 264
 Qy 3472 AACGAGAGGCAAAATGCGCGAAGAAAGGAAATAGGGCGACAGGAAATGTTGAATACT 3531

Db 263 AACGAGAGGCAAAATGCGCGAAGAAAGGAAATAGGGCGACAGCAAGATGTGAATACT 204
 Qy 3532 CATACTCTCTCTTTTCAATATTTATGAGATTTATCAGGATTAATGTCTCATAGAGCG 3591
 Db 203 CATACTCTCTCTTTTCAATATTTATGAGATTTATCAGGATTAATGTCTCATAGAGCG 144
 Qy 3592 ATACATATTTGAATGTATTTAGAAAATTAACAAATAGGGGTTCCGGCAGATTTCCCG 3651
 Db 143 ATACATATTTGAATGTATTTAGAAAATTAACAAATAGGGGTTCCGGCAGATTTCCCG 84
 Qy 3652 AAAAGTCCAC 3662
 Db 83 AAAAGTCCAC 73

RESULT 15
 XXU35129 3886 bp DNA circular SYN 26-SEP-1995
 LOCUS Plasmid pBSU141 cloning vector, complete sequence.
 DEFINITION U35129
 ACCESSION U35129
 VERSION U35129.1 GI:984901
 KEYWORDS
 SOURCE Cloning vector pBSU141
 ORGANISM Cloning vector pBSU141
 REFERENCE 1 (bases 1 to 3886)
 AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
 TITLE Improved antibiotic-resistance gene cassettes and omega elements for Escherichia coli vector construction and in vitro deletion/insertion mutagenesis
 JOURNAL Gene 160 (1), 63-67 (1995)
 MEDLINE 95354958
 PUBMED 7628718
 REFERENCE 2 (bases 1 to 3886)
 AUTHORS Hengen,P.N.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
 FEATURES
 source
 1..3886
 /organism="Cloning vector pBSU141"
 /mol_type="genomic DNA"
 /db_xref="taxon:42702"
 /plasmid="pBSU141"
 complement (806..1339)
 /codon_start=1
 /transl_table=1
 /product="gentamicin 3'-acetyltransferase"
 /protein_id="AAC53614.1"
 /db_xref="GI:984902"
 /translation="MLRSSNDVTQQSRPPTKLGSSWGIIRTCRLGPDQVKSRAALDLFRREGDVAATYSOHQPDVDYGNLRSFTFLAALFDEAVAGALAAVVLPEFEP RSBIIYIDLAVSGEHRROGIATLILNKHANALGAVIYVQADYDDPAVALYTKL GIREVWHPIDISSTAT"
 complement (2898..3758)
 /EC_number="3.5.2.6"
 /function="ampicillin resistance"
 /codon_start=1
 /transl_table=1
 /product="beta-lactamase"
 /protein_id="AAC53615.1"
 /db_xref="GI:984903"
 /translation="MSIQHRRVALLIPFAFLCPVFAHPETLVKVKDAEDQLGARVGI IEDLNSGKILSFPRPEPDMSTEFLLCGAVLSRIDAGQQLGRRIRHYSQNDLVE YSPVTEKHLTDGMTVRELCSAAITMSDNTANILTTIGSPKEITLFLHMGHTVTL DRIPELNEAIPIIDERITTPVMAATLRLLTGLIELTLASRQOLDIMEDADKVAQPL LRSALPAGWFIADKSGRGSRGIITAAAGDPDKPSRIIVITYTGTGATWDENRDQIA EIGASLIKHM"

ORIGIN
 Query Match 68.7%; Score 2514.6; DB 12; Length 3886;
 Best Local Similarity 80.8%; Pred. No. 0;
 Matches 3142; Conservative 0; Mismatches 519; Indels 226; Gaps 9;

QY 1 CTAATGTGAGCGCTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC 60
 Db 1 CTAATGTGAGCGCTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC 60
 QY 61 ATTTTAAACCAATAGCGGAAATCGGCAAAATCCCTTAATATCAAAAGATAGACCGA 120
 Db 61 ATTTTAAACCAATAGCGGCAAAATCCCTTAATATCAAAAGATAGACCGA 120
 QY 121 GATAGGCTTGAGCTGTTTTCAGATTGGAACAAGATCCATTTAAAGAACGTGACCTC 180
 Db 121 GATAGGCTTGAGCTGTTTTCAGATTGGAACAAGATCCATTTAAAGAACGTGACCTC 180
 QY 181 CAACGTCAAAAGGCGGAAACCGCTCTATCAGGCGGATGCGCCACTACGTGAACCATCAC 240
 Db 181 CAACGTCAAAAGGCGGAAACCGCTCTATCAGGCGGATGCGCCACTACGTGAACCATCAC 240
 QY 241 CTAATCAAGTTTGTGGGCTGAGCTGCGTAAAGCACTAAATCGAACCTTAAGGAG 300
 Db 241 CTAATCAAGTTTGTGGGCTGAGCTGCGTAAAGCACTAAATCGAACCTTAAGGAG 300
 QY 301 CCCCCGATTAGAGCTTGAAGCGGAAAGCGCGGAAAGCTGCGGAAAGGAGGAG 360
 Db 301 CCCCCGATTAGAGCTTGAAGCGGAAAGCGCGGAAAGCTGCGGAAAGGAGGAG 360
 QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGAAGCGTCAAGCTGCGGTAACAC 420
 Db 361 AGCGAAAGAGCGGCGCTAGGCGCTGCAAGTGAAGCGTCAAGCTGCGGTAACAC 420
 QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCATTCGCAATTCAGGCTGCG 480
 Db 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCATTCGCAATTCAGGCTGCG 480
 QY 481 CAACCTTGGGAAGGCGGATCGCTGCGGCGCTCTTCTGCTATTAAGCGGCTGCGGAAAG 540
 Db 481 CAACCTTGGGAAGGCGGATCGCTGCGGCGCTCTTCTGCTATTAAGCGGCTGCGGAAAG 540
 QY 541 GCGATGCTGCAAGGCGGATTAAGTGGGTAAAGCGGCTTTCAGAGTCAAGACGCTG 600
 Db 541 GCGATGCTGCAAGGCGGATTAAGTGGGTAAAGCGGCTTTCAGAGTCAAGACGCTG 600
 QY 601 TAAACGAGCGGCACTGAGCGCGCTGCTGCTC----- 631
 Db 601 TAAACGAGCGGCACTGAGCGCGCTGCTGCTC----- 631
 QY 631 -----ATTCAAGTTTGAACCCGTGAGGAGCGGCA 663
 Db 631 -----ATTCAAGTTTGAACCCGTGAGGAGCGGCA 663
 QY 661 GCGCGCGCTCGAGGTCAGGTAATCGATTAAGCTTGAATCGAATTCCTCAGCGCGGAG 720
 Db 661 GCGCGCGCTCGAGGTCAGGTAATCGATTAAGCTTGAATCGAATTCCTCAGCGCGGAG 720
 QY 721 ATCACTAGTTCTAGAGCGGCGGCAAGCGGCTGAGCTCGAATTGACCGCTGCGGAG 780
 Db 721 ATCACTAGTTCTAGAGCGGCGGCAAGCGGCTGAGCTCGAATTGACCGCTGCGGAG 780
 QY 781 AGCGGATCTCGGCTTGAAGCAATTTGATGAGTGGGCTTGGGCTTGAATCAAGAGTCA 840
 Db 781 AGCGGATCTCGGCTTGAAGCAATTTGATGAGTGGGCTTGGGCTTGAATCAAGAGTCA 840
 QY 841 TCACTTCTTCCGTATGCCCACTTTGATAGAGGCACTGCGGAGTCAAGCTTGAAT 900
 Db 841 TCACTTCTTCCGTATGCCCACTTTGATAGAGGCACTGCGGAGTCAAGCTTGAAT 900
 QY 880 ACGACTCAATAGGCGGAAATGGGTAACCGGCGC-----CCCCCTGAGAGTCAAGCTTGAAT 960
 Db 880 ACGACTCAATAGGCGGAAATGGGTAACCGGCGC-----CCCCCTGAGAGTCAAGCTTGAAT 960
 QY 961 TGAAGCTTGAATGCAATTCCTGCAAGCGGCG----- 908
 Db 961 TGAAGCTTGAATGCAATTCCTGCAAGCGGCG----- 908
 QY 909 -----GAGT 912
 Db 909 -----GAGT 912
 QY 1021 AGATCTCACTAGCGGCTGCTCAAACTGGGCAAGACGTAAAGCGGAGAGCGCAACGA 1080
 Db 1021 AGATCTCACTAGCGGCTGCTCAAACTGGGCAAGACGTAAAGCGGAGAGCGCAACGA 1080

QY 913 CCACTAGTCTTAGAGCGGCGGCAACCGCGGTGAGCTCGAGCTTTTGTCCCTTAGTGA 972
 Db 1081 CCGCTTCTTGTTCAGAGGCAAGCGGATGATGATCTTACCTAGGAGCAAGTTCGCA 1140
 QY 973 GGGTTAATTAAGTCCATGCGGTCAATTTTACGAGACTATCTTTCAAGGTTAATCTAGC 1032
 Db 1141 GGTATCGGAGTCCGCGTGAATGTTGGAGAGTGGCTAGCTCTCGAATCAAGACCGA 1200
 QY 1033 TGCAATCAAGATCAATATGTCGGGCTTTTTCGCGCTCAATGATCGCCAGCTGCGCT 1092
 Db 1201 AAAGATCAAGAGAGCGCCGATGATGATTAATCTTGTGACGCGGAGCTTACATGAGCA 1260
 QY 1093 ATCTGGGATCGGAGGAGAAAGAGCCGCTGCTTTTCCGCGAGGTTGAAGCGGATGA 1152
 Db 1261 TGAATGCCATATCTTGAAGCAACCTAATCTTTTGAAGGAGTCCCTGCTGCTAATCAT 1320
 QY 1153 AAGAGTT--TGCGAGAGATGACTGCTGCTGCTGCTTGAAGCGAAGAACGAGTTTAC 1210
 Db 1321 CGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
 QY 1211 CATGATGATTCGAGAAAGTGTGGCCATGACAGCGCTTTAAGGTGAATCTTCTTCAAGC 1270
 Db 1380 ACGCGCTTGTCTGATGAGCGGAGATGAGCTGTAACAAAAACAGCTATTAACAAGC 1439
 QY 1271 CACTGGGATTAACAGTTGCTGCGGCTTTTTCGAGCAAGTTCGAGTGTGAGCGGAA 1330
 Db 1440 CATGAAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1499
 QY 1331 GCGCATCAAGCAACCGCAATTAACGCGGAGAGCGGAGTGTGCGGCTGCTGCTGCTGCT 1387
 Db 1500 GCGTGAAGCTTAACGCTATCTGATTAACGCTTGAACGAGCAAGCTTATGCTCAATTC 1559
 QY 1388 GATTATGAAGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1447
 Db 1560 GAGATTAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
 QY 1448 GATCCCGCAAAATGA----- 1464
 Db 1620 GATATCAAAATTCGAGCGCGGAGATCACTAGTTTGAAGCGGCGGCAACCGCGGTG 1679
 QY 1465 -----CATGGAATACCCCGTGAATTAACCGGCGGCGGCTTGGGCTATCAATGCTC 1515
 Db 1680 GAGCTTCAGCTTTTGTTCCTTGAATGAGGTTAAATGCGGCTTGGGCTATCAATGCTC 1739
 QY 1516 ATAGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1575
 Db 1740 ATAGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
 QY 1576 AAGCATTAAGTGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1635
 Db 1800 AAGCATTAAGTGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1859
 QY 1636 GCGCTCACTGCGGCTTTCAGTGGGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1695
 Db 1860 GCGCTCACTGCGGCTTTCAGTGGGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919
 QY 1696 CCAAGCGGCGGAGAGCGGCTTTCAGTGGGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1755
 Db 1920 CCAAGCGGCGGAGAGCGGCTTTCAGTGGGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
 QY 1756 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1815
 Db 1980 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039
 QY 1816 ACGGTTATCCACAAATCAAGGAGTAAAGGAGAAAGAAATGAGTGAAGAAAGGAGCAAG 2099
 Db 2040 ACGGTTATCCACAAATCAAGGAGTAAAGGAGAAAGAAATGAGTGAAGAAAGGAGCAAG 2159
 QY 1876 AAAGGCGAGAGACGTTAAAGGCGGCTTTCAGTGGGCTTTCAGTGGGCTTTCAGTGGGCTTTCAGT 1935
 Db 2100 AAAGGCGAGAGACGTTAAAGGCGGCTTTCAGTGGGCTTTCAGTGGGCTTTCAGTGGGCTTTCAGT 2159
 QY 1936 TGAAGGATCACTCAAAATCAAGCTTCAAGTCAAGGCTGCGAAGCGGAGAGCTATA 1995

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:11:40 (Search time 11509 Seconds
(without alignments)
11594.581 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662
Sequence: 1 ctaaatgtgaagcgtatata.....attccccgaagaatgcccac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055.8	28.8	1070	1	AJ281552 4A3A-PeF1
2	986.4	26.9	1013	4	BM438846 IplvT0015
3	955.2	26.1	1067	1	AU081137 AU081137
4	924	25.2	1004	1	AJ281480 4A3A-P4G8
5	923.8	25.2	1089	1	AU081124 AU081124
6	902.8	24.7	1049	9	CU021189 CH216-8A1
7	889.8	24.3	928	7	CO487414 GQ0227.B7
8	881.4	24.1	925	6	CB886151 Bn01b_020
9	870.2	23.8	935	4	BG838279 GC01_10e0
10	865	23.6	1073	6	CF269652 Fcy1C01d8
11	857.6	23.4	885	1	CU076016 CH216-138
12	841	23.0	841	1	AL042026 DXF2p434E
13	824.6	22.5	854	4	BM438950 IplvT0049
14	824	22.5	1169	9	AG332951 Mus muscu
15	821.4	22.4	1025	9	CU021193 CH216-8A1
16	820.2	22.4	865	7	CK125894 BES182411
17	819.6	22.4	1249	8	BZ572284 msh2_2572
18	814.4	22.2	1126	8	BZ577702 msh2_5533
19	809.4	22.1	856	7	CN823189 Oa.sp1bn
20	808	22.1	819	6	CD645375 Cvdgdn008
21	807.6	22.1	966	8	BZ570738 msh2_1513
22	797.6	21.8	1011	8	BZ576726 msh2_5071
23	793	21.7	827	7	CN823902 Oa.sp1bn
24	788.6	21.5	833	6	CB686421 Bn01b_04T

25	786.8	21.5	1163	1	AU081044 AU081044
26	785.4	21.4	1574	8	BZ572566 msh2_2633
27	777.8	21.2	800	1	AJ281449 4A3A-P4D5
28	774.8	21.2	1336	6	BZ575810 msh2_4637
29	774.4	21.1	789	6	CD280920 CD280920
30	772.4	21.1	954	1	AL044364 DXF2p434C
31	771.4	21.1	832	7	CN822433 Oa.sp1bn
32	767.6	21.0	863	7	CF752100 TGD89 Hum
33	767.6	21.0	1370	8	BZ571721 msh2_2025
34	763.2	20.8	780	5	BQ825693 BQ825693
35	757	20.7	759	6	CD279661 CD279661
36	755.4	20.6	1463	8	BZ571475 BZ571475
37	752.4	20.5	1369	8	BZ579291 BZ579291
38	751.4	20.5	786	7	CN822164 Oa.sp1bn
39	750.4	20.5	752	6	CD646006 SFE0526R
40	746.4	20.4	1194	8	BZ578985 msh2_6095
41	745.4	20.4	844	8	BZ574513 BZ574513
42	744.6	20.3	778	5	BP539067 BP539067
43	743.6	20.3	832	4	BG923768 BG923768
44	741.4	20.2	758	9	CL422788 CL422788
45	741	20.2	741	6	CD279174 CD279174

ALIGNMENTS

RESULT 1
AJ281552 1070 bp mRNA linear EST 30-JUN-2000
4A3A-PeF1-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION
gambiae cDNA clone 4A3A-PeF1, mRNA sequence.
AJ281552
VERSION
AJ281552.1 GI:6929432

KEYWORDS
SOURCE
ORGANISM
Anopheles gambiae (African malaria mosquito)
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.

REFERENCE

1 (bases 1 to 1070)
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Bense, V., Bork, P., Ansong, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

TITLE

CONTACT: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.

JOURNAL
MEDLINE
PUBMED
10841561
Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.

FEATURES

source

1. 1070
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-PeF1"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match 28.8%; Score 1055.8; DB 1; Length 1070;
 Best Local Similarity 99.7%; Pred. No. 3.8e-299;
 Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1904 TTGCTGCGGTTTTTTCATAGAGCTCCGCCCTGACGAGCATACAAAATGACGCTCA 1963
 DB 1 TTGCTGCGGTTTTTTCATAGAGCTCCGCCCTGACGAGCATACAAAATGACGCTCA 60
 QY 1964 AGTCAGAGGTGGGAAACCCGACAGACTATTAAGATACAGGCGTTTCCCTGGAGC 2023
 DB 61 AGTCAGAGGTGGGAAACCCGACAGACTATTAAGATACAGGCGTTTCCCTGGAGC 120
 QY 2024 TCCCTGCGGCTCTCTCTGTCGACCCCTGCGGCTTACCGGATACCTGTCGCTTCTC 2083
 DB 121 TCCCTGCGGCTCTCTCTGTCGACCCCTGCGGCTTACCGGATACCTGTCGCTTCTC 180
 QY 2084 CCTTCGGGAAGCTGGCGCTTTTCTCATAGCTACCGTGTAGTATCTCAGTTCCGTTG 2143
 DB 181 CCTTCGGGAAGCTGGCGCTTTTCTCATAGCTACCGTGTAGTATCTCAGTTCCGTTG 240
 QY 2144 GTGCTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGTTCAAGCCGCTGCGCC 2203
 DB 241 GTGCTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGTTCAAGCCGCTGCGCC 300
 QY 2204 TTATCCGGTAACTATCTCTTTGAGTTCACCCGGTATAGACAAGTTATCCCACTGGCA 2263
 DB 301 TTATCCGGTAACTATCTCTTTGAGTTCACCCGGTATAGACAAGTTATCCCACTGGCA 360
 QY 2264 GCAGCCACTGTGAACAGATTAGCAGACGAGTATGTAGGCGGTGCTACAGAGTTCTTG 2323
 DB 361 GCAGCCACTGTGAACAGATTAGCAGACGAGTATGTAGGCGGTGCTACAGAGTTCTTG 420
 QY 2324 AAGTGTGGCTTAACCTACGCGCTACACTGAAGACAGATTGGTATCTGCGCTGCTG 2383
 DB 421 AAGTGTGGCTTAACCTACGCGCTACACTGAAGACAGATTGGTATCTGCGCTGCTG 480
 QY 2384 AAGCAGTTACCTTCGGAAGAAAGGTGTGCTCTTGTATCCGCAAAACACCGCGCT 2443
 DB 481 AAGCAGTTACCTTCGGAAGAAAGGTGTGCTCTTGTATCCGCAAAACACCGCGCT 540
 QY 2444 GGTAGCGGTGTTTTTTTTTTTGTGACAGCAGATTACGCGCAGAAAAGATCTCA 2503
 DB 541 GGTAGCGGTGTTTTTTTTTTTGTGACAGCAGATTACGCGCAGAAAAGATCTCA 600
 QY 2504 GAAGATCCTTTGATCTTTTTCACGCGGCTGTGAGCTCAGTGAAGCAAACTCAGCTTA 2563
 DB 601 GAAGATCCTTTGATCTTTTTCACGCGGCTGTGAGCTCAGTGAAGCAAACTCAGCTTA 660
 QY 2564 GGGATTTTGTGATGAGATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATA 2623
 DB 661 GGGATTTTGTGATGAGATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATA 720
 QY 2624 TGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGTGTGACAGTTACCAATGC 2683
 DB 721 TGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGTGTGACAGTTACCAATGC 780
 QY 2684 TTATCAGTAGGACCACTATCTCAGGATCTGTCTATTTCTGATCCATATGTTCCCTGA 2743
 DB 781 TTATCAGTAGGACCACTATCTCAGGATCTGTCTATTTCTGATCCATATGTTCCCTGA 840
 QY 2744 CTCCCGGTGCTGATTAATCAATGAGTACGGAAGGCTTACATCTGAGCCCAAGTGTGCA 2803
 DB 841 CTCCCGGTGCTGATTAATCAATGAGTACGGAAGGCTTACATCTGAGCCCAAGTGTGCA 900
 QY 2804 ATGATATCCGAGACCCGAGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCAGACC 2863
 DB 901 ATGATATCCGAGACCCGAGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCAGACC 960
 QY 2864 GGAAGGCGCGAGCGAGAGGTGCTCTGCACTTATCCGCTCATCAGTCAATTAAT 2923
 DB 961 GGAAGGCGCGAGCGAGAGGTGCTCTGCACTTATCCGCTCATCAGTCAATTAAT 1019
 QY 2924 TGTTCGCGGGAAGCTAAGATAGTTCGCGCAGTTAATGTTGGCGAAC 2974

DB 1020 TGTTCGCGGGAAGCTAGATAGTATGTTCCGACGTTAATATGTTGGCGAAC 1070

RESULT 2
 BM438846/c 1013 bp mRNA linear EST 31-JAN-2002
 LOCUS Ipliv00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
 DEFINITION
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 location/Qualifiers

ORIGIN
 Query Match 26.9%; Score 986.4; DB 4; Length 1013;
 Best Local Similarity 99.5%; Pred. No. 1.1e-278;
 Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 2477 ATTACGGCAGAAAAAGGATCTCAAGAGATCTTGTATCTTTCTACGGGCTGTAC 2536
 DB 1013 ATTACGGCAGAAAAAGGATCTCAAGAGATCTTGTATCTTTCTACGGGCTGTAC 954
 QY 2537 GCTCAGTGAACGAAAACTCAGCTTAAGGATTTTGTGATGAGATTATCAAAAAGATC 2596
 DB 953 GCTCAGTGAACGAAAACTCAGCTTAAGGATTTTGTGATGAGATTATCAAAAAGATC 894
 QY 2597 TTCACCTAGATCTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATAG 2656
 DB 893 TTCACCTAGATCTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATAG 834
 QY 2657 TAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTATCTCAGCGATG 2716
 DB 833 TAAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTATCTCAGCGATG 774
 QY 2717 CTATTTGTTCAATCCATGTTGCTCGATCCCGGCTGTGTAGTATCTACGATACGGAG 2776
 DB 773 CTATTTGTTCAATCCATGTTGCTCGATCCCGGCTGTGTAGTATCTACGATACGGAG 714
 QY 2777 GGGTTACATCTTGCCCGCAGTGTGCAATGATTAACGAGACCAAGCTCACCGGCTCCA 2836
 DB 713 GGGTTACATCTTGCCCGCAGTGTGCAATGATTAACGAGACCAAGCTCACCGGCTCCA 654
 QY 2837 GATTTATCAGCAATTAACACGACCGGAGAGGCGGAGCGCAGAAAGTGTCTGCAACT 2896
 DB 653 GATTTATCAGCAATTAACACGACCGGAGAGGCGGAGCGCAGAAAGTGTCTGCAACT 594


```

Oy 2897 TTATCCGCTCCATCCAGTCTATTATTTGTTCCGGGAGAGCTAGAGTAAGTATGTCGCA 2956
Db 593 TTATCCGCTCCATCCAGTCTATTATTTGTTCCGGGAGAGCTAGAGTAAGTATGTCGCA 534
Oy 2957 GTTAATAGTTTGGCCAGAGTGTGGCCATTGCTACAGGACATGAGTGTCAAGCTCGCG 3016
Db 533 GTTAATAGTTTGGCCAGAGTGTGGCCATTGCTACAGGACATGAGTGTCAAGCTCGCG 474
Oy 3017 TTTGGTATGCTTCAATTCAGCTCCGGTTCCCAAGCATCAAGCGGAGTTACA-TGATCCCC 3075
Db 473 TTTGGTATGCTTCAATTCAGCTCCGGTTCCCAAGCATCAAGCGGAGTTACAATGATCCCC 414
Oy 3076 CATGTGCGAAAAAGCGGTAGCTCTTGCTGCTCCGATCCGATCCGTTGCAAGAAAGTAT 3135
Db 413 CATGTGCGAAAAAGCGGTAGCT-CTTGCTGCTCCGATCCGTTGCAAGAAAGTAT 355
Oy 3136 GGCCGAGTGTATCACTCATGTATGAGCAGACTGCATATTTCTTACTGTATGCC 3195
Db 354 GGCCGAGTGTATCACTCATGTATGAGCAGACTGCATATTTCTTACTGTATGCC 295
Oy 3196 ATCCGTAAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 3255
Db 294 ATCCGTAAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 235
Oy 3256 TATGCGGCGACGAGTGTCTGCTGCGCGCTCAATATGCGGATTAATACCGGCGCACATAG 3315
Db 234 TATGCGGCGACGAGTGTCTGCTGCGCGCTCAATATGCGGATTAATACCGGCGCACATAG 175
Oy 3316 CAGAACTTTAAAGATGCTCATCATTTGAAAAAGTTCTTCCGGGCGAAAACTCTCAAGAT 3375
Db 174 CAGAACTTTAAAGATGCTCATCATTTGAAAAAGTTCTTCCGGGCGAAAACTCTCAAGAT 115
Oy 3376 CTAAACCGCTGTGAGATCCAGTTCGATTAACCACTGTCGACCCCACTGATCTTGACG 3435
Db 114 CTAAACCGCTGTGAGATCCAGTTCGATTAACCACTGTCGACCCCACTGATCTTGACG 55
Oy 3436 ATCTTTAACTTTCCACGACGCTTCTGCGTGAGCAAAAAACAGAAAGGCAAAATGC 3489
Db 54 ATCTTTAACTTTCCACGACGCTTCTGCGTGAGCAAAAAACAGAAAGGCAAAATGC 1

RESULT 3
LOCUS AU081137 1067 bp mRNA linear EST 30-JUL-2002
DEFINITION AU081137 Oncorhynchus mykiss Kidney infected by infectious
hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone K12,
mRNA sequence.
ACCESSION AU081137
VERSION AU081137.1 GI:6431485
SOURCE
ORGANISM Oncorhynchus mykiss (rainbow trout)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Proteanchopeterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 1067)
Kono, T., Sakai, M. and Lapetra, S. B.
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
Mar. Biotechnol. 2 (5), 493-498 (2001)
Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES
SOURCE location/Qualifiers
1..1067
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="K12"

```

```

ORIGIN
Query Match 26.1%; Score 955.2; DB 1; Length 1067;
Best Local Similarity 99.0%; Pred. No. 1.7e-269;
Matches 982; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

/tissue_type="kidney"
/clone_id="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/note="common name:rainbow trout ; infected by infectious
hematopoietic necrosis virus"

1494 CGCCCTTGCGCTATATCATATGATAGCTATGCTGTTCTCGTGTGAATTTGTTATCCGCTCAAA 1553
78 CGAGCTTGCGCTATATCATATGATAGCTGTTCTCGTGTGAATTTGTTATCCGCTCAAA 137
1554 TTCCACACATATGAGCGAGCGGAGCATTAAGTAAAGCTGGGGTGCCTTAATGATGA 1613
138 TTCCACACATATGAGCGAGCGGAGCATTAAGTAAAGCTGGGGTGCCTTAATGATGA 197
1614 GCTAATCATATTAATGCTGCTGCTGCTCACTGCCCCCTTTCCAGTCCGGAACCTGTGTGT 1673
198 GCTAATCATATTAATGCTGCTGCTGCTCACTGCCCCCTTTCCAGTCCGGAACCTGTGTGT 257
1674 GCCAGCTGCATTAATGATCGGCCACAGCGCGGAGAGGCGTTTGCGTATTGGGCGCT 1733
258 GCCAGCTGCATTAATGATCGGCCACAGCGCGGAGAGGCGTTTGCGTATTGGGCGCT 317
1734 CTTCGCGCTTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793
318 CTTCGCGCTTCGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
1794 CAGCTCACTCAAAAGCGGTATATGCTTATTCACAGAAATCAGGGGATTAACGCAAGAAAGA 1853
378 CAGCTCACTCAAAAGCGGTATATGCTTATTCACAGAAATCAGGGGATTAACGCAAGAAAGA 437
1854 ACATGTGAGCAAAAGCGAGCAAAAGCGAGAAACCGTAAAGACCGCGCTGCTGCGCT 1913
438 ACATGTGAGCAAAAGCGAGCAAAAGCGAGAAACCGTAAAGACCGCGCTGCTGCGCT 497
1914 TTTTCATATAGCTCCGCCGCCCTGACAGCATCAAAATTCGACGCTCAAGTCAAGT 1973
498 VTTTCATATAGCTCCGCCGCCCTGACAGCATCAAAATTCGACGCTCAAGTCAAGT 557
1974 GGCGAAACCGGACAGCATTAATTAAGATCAAGGCGTTTCCCGCTGAAAGCTCCCTCGTGC 2033
558 GGCGAAACCGGACAGCATTAATTAAGATCAAGGCGTTTCCCGCTGAAAGCTCCCTCGTGC 617
2034 GCTCTCGCTTCCGACCGCTGCGCTTACCGGATACCTGTGCGGCTTCTCCCTCGGAAA 2093
618 GCTCTCGCTTCCGACCGCTGCGCTTACCGGATACCTGTGCGGCTTCTCCCTCGGAAA 677
2094 GCGTGGCGCTTTCATATGCTCAAGCTATGATGATTCAGTTCGATGAGTTCGTTGCT 2153
678 GCGTGGCGCTTTCATATGCTCAAGCTATGATGATTCAGTTCGATGAGTTCGTTGCT 737
2154 CCAAGCTGGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGCTA 2213
738 CCAAGCTGGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGCTA 797
2214 ACTATGCTTGTAGTCAACCCCGGTAGACACGCTTATGCGCACTGCGACGACGACCTG 2273
798 ACTATGCTTGTAGTCAACCCCGGTAGACACGCTTATGCGCACTGCGACGACGACCTG 857
2274 GTAAACAGATTATGAGAGCGAGGATATGAGCGGTGCTAAGAGTTCTTGAAGTGTGCGC 2333
858 GTAAACAGATTATGAGAGCGAGGATATGAGCGGTGCTAAGAGTTCTTGAAGTGTGCGC 917
2334 CTAACTACGCTTACATAGAAAGACAGTATTTGATATCGGCTCTCTGAAAGCAAGTTA 2393
918 CTAACTACGCTTACATAGAAAGACAGTATTTGATATCGGCTCTCTGAAAGCAAGTTA 977
2394 CCTTGAGAAAGAGTGTAGCTTGTATCCGCAACCAACGAGCTGTGAGCGGT 2453

```

Db 978 CTTTCGAAAAAGTTGTAGCTTTCATCCGCAAAACA-CAACGCTGTAGCGTGT 1036

QY 2454 GTTTTGTTCGACGACGACATTAACGCGC 2485

Db 1037 TTTTGTTCGTC-AGCAGCAGATTAACGCGC 1067

RESULT 4

AJ281480 1004 bp mRNA linear EST 30-JUN-2000

LOCUS 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles

DEFINITION gambiae cDNA clone 4A3A-P4G8, mRNA sequence.

ACCESSION AJ281480

VERSION AJ281480.1 GI:6929360

KEYWORDS EST.

SOURCE Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae

REFERENCE Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

AUTHORS Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoerge, W., Soares, M.B. and Kafatos, F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of generated innate immunity genes from expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

MEDLINE 20300950

PUBMED 10841561

COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1..1004
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4G8"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pTZ19D-pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match 25.2%; Score 924; DB 1; Length 1004;

Best Local Similarity 99.0%; Pred. No. 2,7e-260;

Matches 1001; Conservative 1; Mismatches 2; Indels 7; Gaps 7;

QY 1875 AAAAGCCAGAAACCGTAAAGAGCGCGTGTGCGCTTTTCCATAGCGTCCGCC 1934

Db 1 AAAAGCCAGAAACCGTAAAGAGCGCGCTGTGCGCTTTTCCATAGCGTCCGCC 60

QY 1935 CTGACAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCGACAGACTAT 1994

Db 61 CTGACAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCGACAGACTAT 120

QY 1995 AAAAGTACAGGCGTTTCCCGCTGAGAGCTCCCTGTGCGCTCTCTGTTCCGACCTGC 2054

Db 121 AAAAGTACAGGCGTTTCCCGCTGAGAGCTCCCTGTGCGCTCTCTGTTCCGACCTGC 180

QY 2055 CGCTTACCGGATACCTGTCCGCTTCTCCCTTCGAGAAAGCGTGGCGCTTTCATAGCT 2114

Db 181 CGCTTACCGGATACCTGTCCGCTTCTCCCTTCGAGAAAGCGTGGCGCTTTCATAGCT 240

QY 2115 CACGCTGAGTATCTCAGTTCCGTTGAGGTGGTTGCTCCAGAGCTGGCGTGTGCAGC 2174

Db 241 CACGCTGAGTATCTCAGTTCCGTTGAGGTGGTTGCTCCAGAGCTGGCGTGTGCAGC 300

QY 2175 AACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGTAATCTATCGTTGAGTCCAC 2234

Db 301 AACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGTAATCTATCGTTGAGTCCAC 360

QY 2235 CGGTAAAGACAGACTTATCGCCACGTGGAGCGCCATCGTAAACAGAGATTAACAGCGA 2294

Db 361 CGGTAAAGACAGACTTATCGCCACGTGGAGCGCCATCGTAAACAGAGATTAACAGCGA 420

QY 2295 GGTATGTAGCGCGTGTCTCAAGATTTTGAAGTGTGCTTAACATCAAGCTTACACTA 2354

Db 421 GGTATGTAGCGCGTGTCTCAAGATTTTGAAGTGTGCTTAACTAAGCTTACACTA 480

QY 2355 GGAAGTATTTGGTATCTGCGCTGTGAGACCGCTTACCTTGGAAAAAGATTGTA 2414

Db 481 GGAAGTATTTGGTATCTGCGCTGTGAGACCGCTTACCTTGGAAAAAGATTGTA 540

QY 2415 GCTCTTGATCCGCAAAACAACACCGCTGTAGCGGTGTTTGTGTTGCAAGACG 2474

Db 541 GCTCTTGATCCGCAAAACAACACCGCTGTAGCGGTGTTTGTGTTGCAAGACG 600

QY 2475 AGATTACGGCGAGAAAAAGATCTCAAGAGATCTTTGATCTTTTCTAACGGGCTCG 2534

Db 601 AGATTACGGCGAGAAAAAGATCTCAAGAGATCTTTGATCTTTTCTAACGGGCTCG 660

QY 2535 ACCGTCAGTGGAAACAAATCAAGTTAAGGATTTTGTGATGATTAACAAAGGA 2594

Db 661 ACCGTCAGTGGAAACAAATCAAGTTAAGGATTTTGTGATGATTAACAAAGGA 720

QY 2595 TCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAATCA 2654

Db 721 TCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAATCA 780

QY 2655 AGTAACTGTGTCTGACAGTTACCAATGCTTAATCAAGTGGAGCACTATCTCGCATCT 2714

Db 781 AGTAACTGTGTCTGACAGTTACCAATGCTTAATCAAGTGGAGCACTATCTCGCATCT 839

QY 2715 GTCTATTGTTGTTATCCATGTTGCTGACTCCCGCTGTGATGATTAACGATACGCG 2774

Db 840 GTCTATTGTTGTTATCCATGTTGCTGACTCCCGCTGTGATGATTAACGATACGCG 898

QY 2775 AGGCTTACCATCTGCGCCAGTGTGCAATGATTAACGAGAACCAAGCTACCGGCTC 2834

Db 899 AGGCTTACCATCTGCGCCAGTGTGCAATGATTAACGAGAACCAAGCTACCGGCTC 955

QY 2835 CAGATTATCAGAAATTAACCAAGCCAGCGGAGAGCGGAGGAGGAGT 2885

Db 956 CAGATTATCAGAAATTAACCAAGCCAGCGGAGGAGGAGGAGGAGT 1004

RESULT 5

AU081124 1089 bp mRNA linear EST 30-JUL-2002

LOCUS AU081124

DEFINITION hematopoietic necrosis mykiss Kidney infected by infectious

ACCESSION AU081124.1 GI:6431472

VERSION AU081124

KEYWORDS EST.

SOURCE Oncorhynchus mykiss (rainbow trout)

ORGANISM Oncorhynchus mykiss

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

AUTHORS Kono, T., Sakai, M. and Iapetra, S.E.

TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus

Db	215	ATGATGTAGCTTAATCTCAATTAATTTGGTGTGGCTGACTGCCCCGCTTTCAGTCGGGAAA	274
QY	1666	CTGTCTGTGCCAGCTGCATTAATGATCGGACCAACGGCGGGAGAGGGGTTCGAT	1725
Db	275	CCTGTCTGTGCCAGCTGCATTAATGATCGGACCAACGGCGGGAGAGGGGTTCGAT	334
QY	1726	TGGGCGCTCTTCGCGCTTCTCGCTCACTGACTGCTGCGCTGCTGCTGCTGCGCTG	1785
Db	335	TGGGCGCTCTTCGCGCTTCTCGCTCACTGACTGCTGCGCTGCTGCTGCTGCGCTG	394
QY	1786	AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTATCCACAGATCAGGGGATAAGC	1845
Db	395	AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTATCCACAGATCAGGGGATAAGC	454
QY	1846	AGGAAAGAAATGTGAGCAGAAAGCCAGCAAGAAAGCCGTAAGAAAGCCGCGCTT	1905
Db	455	AGGAAAGAAATGTGAGCAGAAAGCCAGCAAGAAAGCCGTAAGAAAGCCGCGCTT	514
QY	1906	GCTGCGCTTTTTCATAGGCTCCGCCCTCAGCAGCATCAGAAATTCAGCGCTCAG	1965
Db	515	GCTGCGCTTTTTCATAGGCTCCGCCCTCAGCAGCATCAGAAATTCAGCGCTCAG	574
QY	1966	TCAGAGGTGCGAAACCCGACAGACTAATAAGATACAGAGCGTTTCCCTGGAAGCTC	2025
Db	575	TCAGAGGTGCGAAACCCGACAGACTAATAAGATACAGAGCGTTTCCCTGGAAGCTC	634
QY	2026	CTCTGTGCGCTCTCTGTTTCGACCCCTGCGCTTACCGATACCTGTCCGCTTTCGCC	2085
Db	635	CTCTGTGCGCTCTCTGTTTCGACCCCTGCGCTTACCGATACCTGTCCGCTTTCGCC	694
QY	2086	TTGCGGAAACGTGCGCTTTCATAGCTCAGCGTAGAGATATTCAGTTGCGTGAAGT	2145
Db	695	TTGCGGAAACGTGCGCTTTCATAGCTCAGCGTAGAGATATTCAGTTGCGTGAAGT	754
QY	2146	CGTTGCTCAAGCTGGGCTGTGTCAGAAACCCCGGTTACAGCCACGCTGCGCTT	2205
Db	755	CGTTGCTCAAGCTGGGCTGTGTCAGAAACCCCGGTTACAGCCACGCTGCGCTT	814
QY	2206	ATCCGGTAACTATGCTCTTGTAGTCCAAACCGGGTAGACAGACTTATCCGCACTGCGACG	2265
Db	815	ATCCGGTAACTATGCTCTTGTAGTCCAAACCGGGTAGACAGACTTATCCGCACTGCGACG	874
QY	2266	AGCCACTGTGTAAACAGGATTAGCAAGCCAGGATATGTAAGGCGGTGCTACAGAGTCTTGA	2325
Db	875	AGCCACTGTGTAAACAGGATTAGCAAGCCAGGATATGTAAGGCGGTGCTACAGAGTCTTGA	934
QY	2326	GTGTGTGCTTAACAGGCTACGCTACGTACAGAGAC-AGATTTGGTATCTGCGCTGTGCTGA	2384
Db	935	GTGTGTGCTTAACAGGCTACGCTACGTACAGAGAC-AGATTTGGTATCTGCGCTGTGCTGA	994
QY	2385	AGCGAGTTAC-TTTCGGAAGAAAGATTGTAGTCTTGTATCCGGAACAAAC 2437	
Db	995	AACCAATTACTTTTCGGAAGAAAGATTGTAGTCTTGTATCCGGAAGAAAC 1048	
RESULT 7			
LOCUS	CO487414	928 bp	mRNA
DEFINITION	GM0227.B7.1 K04 GM022: ROOT XYLEM - mature trees Picea glauca cDNA		
ACCESSION	CO487414		
VERSION	CO487414.1	GI:50174316	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL
COMMENT

Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, Canada G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
University of Minnesota, MN id Identifier: MN5286152 Clone ID:
GQ0227.B7.K04 Clones available through: John Mackay, Ph. D.
Professeur adjoint -Assistant professor EMAIL:
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
(Forest Biology Research Center) Universite Laval Quebec, Quebec
CANADA G1K 7P4
Plate: 7.B7 row: 04 column: K
Seq primer: T7 primer
Location/Qualifiers
1. 928
/organism="Picea glauca"
/mol_type="mRNA"
/strain="Three trees Unknown"
/db_xref="taxon:33310"
/clone="GQ0227.B7.K04"
/sex="Hermaphrodite"
/tissue_type="Differentiating xylem from roots 1 cm in
diameter or larger"
/dev_stage="Non-lignified xylem tissues from larger roots
in early part of growing season (June)."
/lab_host="E. coli DH10B cells"
/clone_1ib="GQ022: ROOT XYLEM - mature trees"
/note="Organ: Roots from 9 year old trees measuring
approximately 4.5 m tall, and 10 cm in diameter. ; Vector:
pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I;
cDNA was prepared from 5 mg of poly A+ selected RNA and
was directionally ligated into the pBluescript II SK (+)
XR vector (Stratagene), transformed by electroporation
into DH10B cells (in vitro) for propagation"

ORIGIN

Query Match	24.3%	Score 889.8	DB 7	Length 928
Best Local Similarity	96.0%	Pred. No. 3.2e-250		
Matches 891	Conservative 0	Mismatches 37	Indels 0	Gaps 0
QY 1915	TTTCATAGAGGCTCCGCCCCCTCGAGACATCCAAATAAGAGCTCAAGTCAGAGSTG	1974		
DB 1	TTTTCATAGAGGCTCCGCCCCCTCGAGACATCCAAATAAGAGCTCAAGTNNNNNG	60		
QY 1975	GCGAAACCCGACAGAGCTATTAAGAATACAGAGCGTTTCCCTGGAAGCTCCCTCGTGGC	2034		
DB 61	GCGAAACCCGACAGAGCTATTAAGAATACAGAGCGTTTCCCTGGAAGCTCCCTCGTGGC	120		
QY 2035	CTCTCCGTTCCGACCCGCGCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAG	2094		
DB 121	CTCTCTGTTCCGACCCGCGCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAG	180		
QY 2095	CGTGGCGTTTCTCATAGCTCAAGCTCTAGGTATCTCAGTTCGGTATGAGTTCGCTC	2154		
DB 181	CGTGGCGTTTCTCATAGCTCAAGCTCTAGGTATCTCAGTTCGGTATGAGTTCGCTC	240		
QY 2155	CAAGCTGGGCGTGTGACAGAACCCCGCTTACGCCGACCGCTGCGGCTTATCCGGTAA	2214		
DB 241	CAAGCTGGGCGTGTGACAGAACCCCGCTTACGCCGACCGCTGCGGCTTATCCGGTAA	300		
QY 2215	CTATCGTCTTGAAGTCCAAACCGGTAAAGACAGACTTATGCGCACCTGGACAGACCACTGG	2274		
DB 301	CTATCGTCTTGAAGTCCAAACCGGTAAAGACAGACTTATGCGCACCTGGACAGACCACTGG	360		
QY 2275	TAAACAGATTAAGACAGACGAGATATGAGCGGTGCTAACAGATTCTTGAAGTGTGGCC	2334		
DB 361	TAAACAGATTAAGACAGACGAGATATGAGCGGTGCTAACAGATTCTTGAAGTGTGGCC	420		
QY 2335	TAACTACGGCTACACTAGAAAGACATATTTGGTATCTCGGCTTGGCTGAAGCCAGTTAC	2394		

Db 421 TAACTACGGCTACACTAGAAAGATTTGTATCTGCGCTGTGTGAAGCCAGTTAC 480
Qy 2295 CTTGGGAAAAAGATTGGTACTCTTGTATCCGGGAAACAAACACCGCTGGTAGGGTGG 2454
Db 481 CTTGGGAAAAAGATTGGTACTCTTGTATCCGGGAAACAAACACCGCTGGTAGGGTGG 540
Qy 2455 TTTTGTGTGTGAAGCAGCAGATTAACGGCAGAAAAAGAGATCTCAAGAGATCTTT 2514
Db 541 TTTTGTGTGTGAAGCAGCAGATTAACGGCAGAAAAAGAGATCTCAAGAGATCTTT 600
Qy 2515 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTGGT 2574
Db 601 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTGGT 660
Qy 2575 CATAGATATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATGAAGTTTAA 2634
Db 661 CATAGATATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATGAAGTTTAA 720
Qy 2635 ATCAATCTAAAGTATATATGATGTAATCTGGTCTGACAGTTACCAATGCTTAATCAGTGA 2694
Db 721 ATCAATCTAAAGTATATATGATGTAATCTGGTCTGACAGTTACCAATGCTTAATCAGTGA 780
Qy 2695 GGCACTATCTCAGCGATCTGTCTATTTCTTCAATCCATAGTTGCTGACTCCCGTGT 2754
Db 781 GGCACTATCTCAGCGATCTGTCTATTTCTTCAATCCATAGTTGCTGACTCCCGTGT 840
Qy 2755 GTAATATCTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACGGC 2814
Db 841 GTNNNTATCTACGATACGGAGGGCTTACCATCTGNNNNNGTGTGCAATGATACNNCG 900
Qy 2815 AGACCCACGCTCACCGGCTCCAGATTAA 2842
Db 901 AGACCCACGCGNNNNNNNNNNNNNTTAA 928

RESULT 8
CB686151/c
LOCUS
DEFINITION
Bn01b_02008_A 925 bp mRNA linear EST 09-APR-2003
Bn01b_AAPC_ECORC_transgenic_Brassica_napus_overexpressing_BNCP17_c
onstitutively_frost_tolerant Brassica napus cDNA clone Bn01b_02008,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
CB686151
CB686151.1 GI:29689876

Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 925)
Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moore, A.,
Chagnon, J., Farah, S., Couroux, P. and Hattori, J.
Expressed Sequence Tags from constitutively frost tolerant
transgenic Brassica napus overexpressing BNCP17
Unpublished (2002)
Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
KW Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
0G6, Canada

Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@agr.gc.ca.

FEATURES
source
Location/Qualifiers

1..925
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Weibull"
/db_xref="taxon:3708"
/clone="Bn01b_02008"
/issue_type="Fourth leaf"
/dev_stage="3 weeks seedling grown at room temperature"

/clone_id="Bn01b AAPC_ECORC_transgenic_Brassica_napus_ove
repressing_BNCP17_constitutively_frost_tolerant"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
Site 2: XhoI; Germinated in soil flats and seedlings grown
for 3 weeks in a Conviron E-15 cabinet set at 20°C/16 hr
light (250 Em-2sec-1) and 16°C/8 hr dark. Fourth leaves
collected at 9 am and immediately frozen."

ORIGIN

Query Match 24.1%; Score 881.4; DB 6; Length 925;
Best Local Similarity 99.0%; Pred. No. 9,76-248;
Matches 890; Conservative 6; Mismatches 2; Indels 1; Gaps 1;
1493 GCGGCTTGGGATATCATGTCTATAGCTGTTTCTGTGTGAATTTGATCCGCTACA 1552
Db 898 GCGGCTTGGGATATCATGTCTATAGCTGTTTCTGTGTGAATTTGATCCGCTACA 840
Qy 1553 ATTCCACAAACAACGAGCGCGAAGATTAAGGTAAAGCTGGGGTGCTTAATGAGTG 1612
Db 839 ATTCCACAAACAACGAGCGCGAAGATTAAGGTAAAGCTGGGGTGCTTAATGAGTG 780
Qy 1613 AGCTAATCAATTATTTGGTGTGCTGCTCACTGCGCTTTCAGTGGGAAACCTGTG 1672
Db 779 AGCTAATCAATTATTTGGTGTGCTGCTCACTGCGCTTTCAGTGGGAAACCTGTG 720
Qy 1673 TGCCAGCTGCATTATGATATGCGCAACGCGCGGGAGAGCGGTTGCTATTGGCGC 1732
Db 719 TGCCAGCTGCATTATGATATGCGCAACGCGCGGGAGAGCGGTTGCTATTGGCGC 660
Qy 1733 TCTTCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1792
Db 659 TCTTCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 1793 TCACTCACTCAAAAGCGGTATATCGGTATCCACAGATTCAGGGGATTAACGAGAAAG 1852
Db 599 TCACTCACTCAAAAGCGGTATATCGGTATCCACAGATTCAGGGGATTAACGAGAAAG 540
Qy 1853 AACATGTGACAAAGGCGCAGCAAAAGCCAGAGAACCTGTAAAGGCGCGTCTGCGC 1912
Db 539 AACATGTGACAAAGGCGCAGCAAAAGCCAGAGAACCTGTAAAGGCGCGTCTGCGC 480
Qy 1913 TTTTTCATATGAGCTCG 1972
Db 479 TTTTTCATATGAGCTCG 420
Qy 1973 TGGGAAACCCGACAGACTATTAAGATATACAGGCGTTTCCCTGGAAGCTCCCTGTG 2032
Db 419 TGGGAAACCCGACAGACTATTAAGATATACAGGCGTTTCCCTGGAAGCTCCCTGTG 360
Qy 2033 CGCTCTCTGTTCGACCGCTGCGCTTACCGGATACCTGTCCGCTTTCCTTGGGGA 2092
Db 359 CGCTCTCTGTTCGACCGCTGCGCTTACCGGATACCTGTCCGCTTTCCTTGGGGA 300
Qy 2093 AGGTGCGCTTCTCTATCTACCTCAGCGGTGAGATCTGAGTTCGAGTTCGAGTTCG 2152
Db 299 AGGTGCGCTTCTCTATCTACCTCAGCGGTGAGATCTGAGTTCGAGTTCGAGTTCG 240
Qy 2153 TCCAGCTGGGCTGTGTGACAGAACCCCGCTTACGCGCAGCGCTGCGCTTACCGGT 2212
Db 239 TCCAGCTGGGCTGTGTGACAGAACCCCGCTTACGCGCAGCGCTGCGCTTACCGGT 180
Qy 2213 AACTATGTTTGAATGCAACCGGTGAAGACAGACTTATGCGCATGCGAGACGCACT 2272
Db 179 AACTATGTTTGAATGCAACCGGTGAAGACAGACTTATGCGCATGCGAGACGCACT 120
Qy 2273 GGTAAACAGATTAGCAGAGGAGGTATGTAGGGGCTTACAGATCTTGAAGTGGTG 2332
Db 119 GGTAAACAGATTAGCAGAGGAGGTATGTAGGGGCTTACAGATCTTGAAGTGGTG 60
Qy 2333 CTTAATACGCGCTTACATGAAGAGAGATTTTGTATCTGCGCTGTGTAAGCCAGT 2391
Db 59 CTTAATACGCGCTTACATGAAGAGAGATTTTGTATCTGCGCTGTGTAAGCCAGT 1

```

RESULT 9
BG838279/c      935 bp      mRNA      linear      EST 25-MAY-2001
LOCUS           GC01_10e07_R GC01_AAPC_ECORC cold stressed_glycine_clandestina
DEFINITION      BG838279
ACCESSION       BG838279
VERSION         BG838279.1 GI:14204601
KEYWORDS        EST.
SOURCE          glycine clandestina
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                Glycine.
REFERENCE       1 (bases 1 to 935)
AUTHORS        Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A.,
                Harris,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprott,D. and
                Tinker,N.A.
TITLE           Expressed Sequence Tags from Cold-Stressed Glycine clandestina
JOURNAL         Unpublished (2001)
COMMENT        Contract: Singh,J.A.
                Eastern Cereal and Oilseed Research Centre
                Agriculture and Agri-food Canada
                KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
                0C6, Canada
                Tel: (613) 759-1662
                Fax: (613) 759-1701
                Email: singhja@agr.gc.ca.
FEATURES
    source
        1..935
        /organism="Glycine clandestina"
        /mol_type="mRNA"
        /culivar="1035"
        /db_xref="taxon:45687"
        /clone="GC01_10e07"
        /cvs_type="leaves, stem"
        /clone_1fb="GC01_AAPC_ECORC_cold_stressed_glycine_clandest
        ina"
        /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
        Site 2: XhoI; Plants incubated at 2 degrees under 12 hours
        of light/day. Harvested after only 2-3 days of cold
        treatment cDNA was prepared with the Uni-Zap cDNA kit
        from Stratagene. Eco RI adapters were linked followed
        by digest with Xho I/Eco RI and ligated to pbluescript."
ORIGIN
Query Match      23.8%; Score 870.2; DB 4; Length 935;
Best Local Similarity 98.1%; Pred. No. 2e-244; Indels 1; Gaps 1;
Matches 874; Conservative 14; Mismatches 2;
QY 2773 GAGAGGCTTACCATCTGCGCCCAAGTGTGCAATGATACGCGAGACCCACGCTCACCGGC 2832
DB 928 GGGGGGGTTTACCATCTGCGCCCAAGTGTGCAATGATACGCGAGACCCACGCTCACCGGC 869
QY 2833 TCCAGATTATTCAGCAATAAACAAGCAGCCGGAAGGCGCGAGCGGAGAAAGTGCTTGC 2892
DB 868 YCCAGATTTATTCAGCAATAAACAAGCAGCCGGAAGGCGCGAGCGGAGAAAGTGCTTGC 809
QY 2893 AACTTATTCGCGCTTCATCCAGTCTATTTATTTGTCGCGGAAGCTAGAGTAAGTATTC 2952
DB 808 AACTTATTCGCGCTTCATCCAGTCTATTTATTTGTCGCGGAAGCTAGAGTAAGTATTC 749
QY 2953 GCCAGTTAATAGTTGGCGGAAGGTTTGTGCAATGCTGTAACGAGCATGTGTGTCAAGCTC 3012
DB 748 GCCAGTTAATAGTTGGCGGAAGGTTTGTGCAATGCTGTAACGAGCATGTGTGTCAAGCTC 689
QY 3013 GTGCTTTGGTATGCTTCAATCAGCTCCGGTTCACCAAGATTAAGCGCAAGTATCATATC 3072
DB 688 GTGCTTTGGTATGCTTCAATCAGCTCCGGTTCACCAAGATTAAGCGCAAGTATCATATC 629
QY 3073 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGTCTCCGATCGTTGTCAAGATGTA 3132

```

```

DB 628 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGTCTCCGATCGTTGTCAAGATGTA 569
QY 3133 GTT-GGGCGGAGTTTATCACTCATGTGTTATGCGACAGTCAATATTCCTTACGTCA 3191
DB 568 GTTGGGCGGAGTTTATCACTCATGTGTTATGCGACAGTCAATATTCCTTACGTCA 509
QY 3192 TGCATCCGTTAAAGATGCTTTTCTGTGATCTGTGATCACTCAACCAAGTATCTGAGAT 3251
DB 508 TGCATCCGTTAAAGATGCTTTTCTGTGATCTGTGATCACTCAACCAAGTATCTGAGAT 449
QY 3252 AGTGTATGCGGCGACGAGTGTCTTGGCCGGCGTCAATAGGGAATATACCGGCCAC 3311
DB 448 AGTGTATGCGGCGACGAGTGTCTTGGCCGGCGTCAATAGGGAATATACCGGCCAC 389
QY 3312 ATAGCAGAACTTTAAAGTCTCATCTATTTGAAAAAGTCTTTCGGGCGGAAACTCTCA 3371
DB 388 ATAGCAGAACTTTAAAGTCTCATCTATTTGAAAAAGTCTTTCGGGCGGAAACTCTCA 329
QY 3372 GGATCTTACCGGCTGTGAGATCCAGTTGATGTAATCCACTGTGCAACCCAACTGATCTT 3431
DB 328 GGATCTTACCGGCTGTGAGATCCAGTTGATGTAATCCACTGTGCAACCCAACTGATCTT 269
QY 3432 CAGCATCTTTTACTTTACACAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCG 3491
DB 268 CAGCATCTTTTACTTTACACAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCG 209
QY 3492 CAAAAAAGGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTCTTTTCAAT 3551
DB 208 CAAAAAAGGGAATAGGCGGACACGGAATGTTGAATCTCATCTCTCTTTTCAAT 149
QY 3552 ATTATTTGAAGCATTTATCAGGTTATTTGTCTCATGAGCGGATACATTTGAATGTTT 3611
DB 148 ATTATTTGAAGCATTTATCAGGTTATTTGTCTCATGAGCGGATACATTTGAATGTTT 89
QY 3612 AAAAAAATTAACAATATAGGGGTTCCGGCACATTTCCCGAAAAAGTCCAC 3662
DB 88 AAAAAAATTAACAATATAGGGGTTCCGGCACATTTCCCGAAAAAGTCCAC 38
RESULT 10
CF269652      1073 bp      mRNA      linear      EST 13-AUG-2003
LOCUS           Fcyl1coid844 Fragilariopsis cylindrus SMART cDNA library (Clontech)
DEFINITION      CF269652
ACCESSION       CF269652
VERSION         CF269652.1 GI:33631539
KEYWORDS        EST.
SOURCE          Fragilariopsis cylindrus
ORGANISM        Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
                Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
REFERENCE       1 (bases 1 to 1073)
AUTHORS        Mock,T. and Valentin,K.
TITLE           EST analysis of freezing tolerance in the Antarctic diatom
                Fragilariopsis cylindrus: Detection of numerous cold adaption
                related genes and gene transfer events
JOURNAL         Unpublished (2003)
COMMENT        Contact: Mock T
                Biological Oceanography
                Alfred-Wegener-Institute for Polar and Marine Research
                Am Handels Hafen 12, D-27570 Bremerhaven, Germany
                Tel: +49 471 4831 1893
                Fax: +49 471 4831 1425
                Email: tmock@awi-bremerhaven.de
                Sequence with unknown function
                PCR Primers
                FORWARD: 5' lambdatrip1EX2
                BACKWARD: 3' lambdatrip1EX2
                Seq primer: ctcgggaagcggccatcgtcgtgt.
FEATURES
    source
        1..1073
        /organism="Fragilariopsis cylindrus"

```

/mol_type="mRNA"
 /strain="Antarctic"
 /db_xref="taxon:186039"
 /clone="Antarctic"
 /clone_1b="Fragiliariopsis cylindrus SMART cDNA library
 (Clontech)"
 /note="Vector: pTriplex2; total polyA was used for
 first-strand synthesis with SMART IV oligos and CDS
 III/3 PCR primer. Double strand cDNA synthesis was done by
 LD PCR using the following program: 95°C for 5 min
 denaturation and subsequent 20 cycles at 95°C (2min) and
 68°C (4min). After SEI digestion the cDNA was
 fractionated with CHROMA Spin-400 columns. These cDNAs
 were ligated overnight into pTriplex2 vectors."

ORIGIN

Query Match 23.6%; Score 865; DB 6; Length 1073;
 Best Local Similarity 96.2%; Pred. No. 7e-243;
 Matches 927; Conservative 0; Mismatches 30; Indels 7; Gaps 4;

1631 GCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAACTCTGTCGACCTGCATTATGA 1690
 1 GCTTCTTGCTCACTGCGCGCTTTCCAGTCGGGAACTCTGTCGACCTGCATTATGA 60

1691 ATCGGCCAAGCGCGGGGAGAGCGGTTGGGTATGGGCGCTTCCGCTTCTCGCTC 1750
 61 ATCGGCCAAGCGCGGGGAGAGCGGTTGGGTATGGGCGCTTCCGCTTCTCGCTC 120

1751 ACTGACTGCTGCGCTGCGCTGCTGCGCGCGAGCGGATACAGTCACTCAAAAGCG 1810
 121 ACTGACTGCTGCGCTGCGCTGCTGCGCGAGCGGATACAGTCACTCAAAAGCG 180

1811 GTAATACGCTTATCCAGAAATCAGGGGATTAACGAGAAAGACATGTGAGCAAAAGC 1870
 181 GTAATACGCTTATCCAGAAATCAGGGGATTAACGAGAAAGACATGTGAGCAAAAGC 240

1871 CAGAAAAGCGCAGGAAACCGTAAAGCGCGCTTGGCTGCTGCTTTCATAGGTCGCG 1930
 241 CAGAAAAGCGCAGGAAACCGTAAAGCGCGCTTGGCTGCTGCTTTCATAGGTCGCG 300

1931 CCCCCTGACGAGCATCAAAAATGAGCTCAAGTCAGAGTGGGCAAAACCCGACAGA 1990
 301 CCCCCTGACGAGCATCAAAAATGAGCTCAAGTCAGAGTGGGCAAAACCCGACAGA 360

1991 CTATTAAGATACAGGCGCTTCCCTGGAAGCTCCCTGAGCGCTCTCTGTTCCAGC 2050
 361 CTATTAAGATACAGGCGCTTCCCTGGAAGCTCCCTGAGCGCTCTCTGTTCCAGC 420

2051 CTGCGCTTACCGGATACCTGCTGCGCTTCTCCCTTGGGAAAGGTCGCTTCTAT 2110
 421 CTGCGCTTACCGGATACCTGCTGCGCTTCTCCCTTGGGAAAGGTCGCTTCTAT 480

2111 AGCTCAGCTGATAGTATCTGATGCTGATGAGTCTGCTCCAGACTGGGCTGTGTG 2170
 481 AGCTCAGCTGATAGTATCTGATGCTGATGAGTCTGCTCCAGACTGGGCTGTGTG 540

2171 CACGAACCCCGCTTACGCGCGGCTGCGCTTATCCGTAATATGTTGAGTCC 2230
 541 CACGAACCCCGCTTACGCGCGGCTGCGCTTATCCGTAATATGTTGAGTCC 600

2231 AACCCGTAAGACAGACTTATCGGCACTGAGAGGACACTGTGTAAGATTAAGCA 2290
 601 AACCCGTAAGACAGACTTATCGGCACTGAGAGGACACTGTGTAAGATTAAGCA 660

2291 GCGAGTATGATGAGCGGTGCTACAGAGTCTTGAAGTGTGGGCTTCACTACGCTACA 2348
 661 GCGAGTATGATGAGCGGTGCTACAGAGTCTTGAAGTGTGGGCTTCACTACGCTACA 720

2349 CTAGAAGACA-GTATTTGTTATCT-GCGCTCTGCTGAAGC--AGTTACCTTTCGAAA 2403
 721 CTAGAAGACAAGTATTTGTTATCTGCTGCTGAAGCAGAGTTACCTTTCGAAA 780

2404 AAGAGTTGATGCTTGTATCCGGAACAAACACCGCTGTAGCGGTGTTTTTGT 2463

DB 781 AAGAGTTGATGCTTGTATCCGGAACAAACACCGCTGTAGCGGTGTTTTTGT 840

QY 2464 TTGACAGCAGAGATTAACGCGCAGAAAAAGATCTCAAGAGATCTTTGATCTTTTC 2523

DB 841 TTGACAGCAGAGATTAACGCGCAGAAAAAGATCTCAAGAGATCTTTGATCTTTTC 900

QY 2524 TACGGGCTTCAAGCTCAAGTGAACGAAATCACTGTTAAGGATTTTGTGATGACATT 2583

DB 901 TACGGGCTTCAAGCTCAAGTGAACGAAATCACTGTTAAGGATTTTGTGATGACATT 960

QY 2584 ATCA 2587

DB 961 ATAA 964

RESULT 11

CL076016 885 bp DNA linear GSS 31-DEC-2003
 LOCUS CH216-138F20, RM1.1 CH216 Xenopus tropicalis genomic clone
 DEFINITION CH216-138F20, genomic survey sequence.
 ACCESSION CL076016
 VERSION CL076016.1 GI:40531929
 KEYWORDS GSS.

ORGANISM

Xenopus tropicalis (western clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodidae; Xenopus; Silurana.

REFERENCE

Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

AUTHORS

TITLE

JOURNAL

COMMENT

A physical map of the xenopus tropicalis genome
 Unpublished (2003)
 Contact: Richard K Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.edu
 Insert Length: 175000 Std Error: 0.00
 Seq primer: RM1 TACGACTCACTATAGGAGGA
 Class: BAC ends
 High quality sequence start: 11
 High quality sequence stop: 810.
 Location/Qualifiers

FEATURES

source

1..885
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /strain="Nigerian frog"
 /db_xref="taxon:8364"
 /clone="CH216-138F20"
 /sex="male"
 /cell_line="Stock 248 F7A2, inbred N7"
 /clone_1b="CH216"
 /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
 BAC library"

ORIGIN

Query Match 23.4%; Score 857.6; DB 9; Length 885;
 Best Local Similarity 99.3%; Pred. No. 1e-240;
 Matches 871; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

1535 AATTGTTATCCGCTCACAATTCCACACATACAGAGCCGGAACATTAAGTAAAGCC 1594
 5 AATTGTTATCCGCTCACAATTCCACACATACAGAGCCGGAACATTAAGTAAAGCC 64

1595 TGGGGTCCCTAATGAGGAGTAACTCACTTAATTGGCTTGGCTGCTGCGCTTTTC 1654
 65 TGGGGTCCCTAATGAGGAGTAACTCACTTAATTGGCTTGGCTGCTGCGCTTTTC 124

1655 CAGTCGGGAAACCTGTGCTGACAGCTGATTAATGAATCGGCCAACCGCGGGAGAGGC 1714
 125 CAGTCGGGAAACCTGTGCTGACAGCTGATTAATGAATCGGCCAACCGCGGGAGAGGC 184

QY	3568	TCAGGGTATTTCCTCAGAGCGGATACATATTGAAGTATTGAAAAATTAACAAAT	362
Db	61	TCAGGGTATTTCCTCAGAGCGGATACATATTGAAGTATTGAAAAATTAACAAAT	2
QY	3628	A 3628	
Db	1	A 1	
RESULT 13			
BM438950 854 bp mRNA linear EST 31-JAN-2002			
LOCUS BM438950/c			
DEFINITION IPIV00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA			
sequence.			
ACCESSION BM438950			
VERSION BM438950.1 GI:18460672			
SOURCE EST.			
ORGANISM Ictalurus punctatus (channel catfish)			
Ictalurus punctatus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;			
Ictaluridae; Ictalurus.			
1 (bases 1 to 854)			
REFERENCE Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.			
Transcriptome of channel catfish (Ictalurus punctatus): initial			
analysis of expressed sequence tags from the liver			
Unpublished (2002)			
AUTHORS Contact: Liu ZJ			
TITLE The Fish Molecular Genetics and Biotechnology Laboratory,			
Department of Fisheries and Allied Aquacultures and Program of Cell			
and Molecular Biosciences			
Auburn University			
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA			
Tel.: 334 844 4054			
Fax: 334 844 9208			
Email: zliu@ceeag.auburn.edu			
Seq primer: M13 Reverse.			
FEATURES			
Source Location/Qualifiers			
1..854			
/organism="Ictalurus punctatus"			
/mol_type="mRNA"			
/db_xref="taxon:7998"			
/clone_lib="Liver cDNA library"			
/note="Organ: Liver; Vector: pSPORT1; Site_1: NotI;			
Site_2: SalI"			
ORIGIN			
Query Match 22.5%; Score 824.6; DB 4; Length 854;			
Best Local Similarity 99.3%; Pred. No.5.4e-231;			
Matches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;			
QY	2636	TCATCTTAAAGTATTATGAGTAAACCTTGCTGCACAGTTACCAATGCTTAATCAGTAC	2695
Db	854	TCAACTCAACAGATATATGAGTAAACCTTGCTGCACAGTTACCAATGCTTAATCAGTAC	795
QY	2696	GCACCTATCTCAGCGCATCTGTCTATTGGTTCATCAATGTTGCTGCTCCCGTGTG	2755
Db	794	GCACTTATCTCAGCGCATCTGTCTATTGGTTCATCAATGTTGCTGCTCCCGTGTG	735
QY	2756	TAGATTAACATACGATACGGAGGGCTTACATCTGGCCCCAGTGTGCAATGATACCGGA	2815
Db	734	TAGATTAACATACGATACGGAGGGCTTACATCTGGCCCCAGTGTGCAATGATACCGGA	675
QY	2816	GACCCAGCCTCAACGGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAG	2875
Db	674	GACCCAGCCTCAACGGGCTCCAGATTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAG	615
QY	2876	CGCAGAAATGATCCGCAACTTATCCGCCTCCATCCAGTCAGTCAATTAATTGTTGCCGGGA	2935
Db	614	CGCAGAAATGATCCGCAACTTATCCGCCTCCATCCAGTCAGTCAATTAATTGTTGCCGGGA	555
QY	2936	GCTAGAGTAAGTAGTTGCCAGATTAATAGTTGGCGAAGTGTTGCCATTGCTACAGGC	2995

D	b		554	GCTAGAGTAATGATTGGCCAGTTAAATAAGTTTGGCAACGTTGTTCGACTTGCAACAGGC	495
O	y		2996	ATCGTGATGTCACGCTGCTGCTTTGGTAATGGCTTCATTCAAGCTCCGGTCCCAAGATCA	3055
D	b		494	ATCGGAGTGTCACGCTCCTCGTTTGGTAATGGCTTCATTCAAGCTCCGGTCCCAAGATCA	435
O	y		3056	AGGCGAGTTAATGATGCCCATGCTGTCGAAAAAAGCGGTTAGCTCCTTCGGCTCCG	3115
D	b		434	AGGCGAGTTAATGATGCCCATGCTGTCGAAAAAAGCGGTTAGCTCCTTCGGCTCCG	376
O	y		3116	ATCGTTGTCAGAAGTAAGTTGGCCGCAAGTGTATCACTCATGTATTAGCAGACTGCAT	3175
D	b		375	ATCGTTGTCAGAAGTAAGTTGGCCGCAAGTGTATCACTCATGTATTAGCAGACTGCAT	316
O	y		3176	AATTCTCTTACTGTGTCAGCCCATCCGTAAGATGCTTTCTGTGAATCGGTGATGTAATCAAC	3235
D	b		315	AATTCTCTTACTGTGTCAGCCCATCCGTAAGATGCTTTCTGTGAATGATGTAATCAAC	256
O	y		3236	AAGTCATTCTGAGATAGTGTATCGCGCAGCCGAGTGTCTTGGCCCGGCGTCAATACG	3295
D	b		255	AAGTCATTCTGAGATAGTGTATCGCGCAGCCGAGTGTCTTGGCCCGGCGTCAATACG	196
O	y		3286	GATTAATACCGCGCCCATGACGACAATTAAAAGTCTCATCATTTGGAAAAAGCTTCTTCG	3355
D	b		195	GATTAATACCGCGCCCATGACGACAATTAAAAGTCTCATCATTTGGAAAAAGCTTCTTCG	136
O	y		3356	GGGCGAAATCTCTAAGGATCTTACCGCTGTGATGATCAAGTGCATGTA-CCCACTCG	3414
D	b		135	GGGCGAAATCTCTAAGGATCTTACCGCTGTGATGATGATGATGATGTAAGCCAATCG	76
O	y		3415	TGACCCCACTGATCTTACGATCTTTTACTTTTACCAGCGTTTTCTGGGTGAGCAAAAC	3474
D	b		75	TGACCCCACTGATCTTACGATCTTTTACTTTTACCAGCGTATCTGGGTGAGCAAAAC	16
O	y		3475	AGGAAGGCAAAATGC 3489	
D	b		15	AGGAAGGCAAAATGC 1	
<hr/>					
RESULT 14					
AG332951					
LOCUS Mus musculus molossinus DNA, clone:MSNg01-123f02.T7, genomic survey					
DEFINITION sequence.					
ACCESSION AG332951					
VERSION AG332951.1 GI:47906261					
KEYWORDS GSS.					
SOURCE Mus musculus molossinus					
ORGANISM Mus musculus molossinus					
REFERENCE Hattoni,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.					
AUTHORS 1					
TITLE BAC end Sequences of Library MSNg01					
JOURNAL Unpublished					
AUTHORS 2 (bases 1 to 1169)					
TITLE Hattoni,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.					
JOURNAL Direct Submission					
AUTHORS Submitted (17-NOV-2003) Masahira Hattoni, The Institute of Physical					
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC);					
COMMENT 1-7-22 Shutoho-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan					
(E-mail:hattoni@gsc.riken.jp; URL:http://hnp.gsc.riken.go.jp/,					
Tel.:81-45-503-9111, Fax:81-45-503-9170).					
Clones are derived from the mouse BAC library MSNg01. For BAC					
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).					
Tsukuba Institute, Bio Resource Center,					
The Institute of Physical and Chemical Research (RIKEN) 3-1-1					
Koyadai, Tsukuba, 305-0074 Japan					
phone: 81-298-36-9189, fax: 81-298-36-9199					
e-mail: abe@rtc.riken.jp					
PRIMERS					
Sequencing : T7					

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:20:55 ; Search time 1735.03 Seconds
(without alignments)
11079.575 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662
Sequence: 1 ctaaatgttaagcgttaata.....attccccgaagtcgcac 3662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3648.4	99.6	3661	10	ACA55354
2	2665.4	72.8	3637	11	ADM68463
3	2665.4	72.8	3637	12	ADP26616
4	2664.8	72.8	3637	8	ABX14571
5	2664.8	72.8	3637	11	ADM68448
6	2664.8	72.5	3637	12	ADP26601
7	2664.4	72.5	3637	8	ABX14570
8	2664.4	72.5	3637	11	ADM68462
9	2664.4	72.5	3637	12	ADP26615
10	2667.8	72.3	3637	10	ADP26329
11	2635.8	72.0	3681	2	AAQ13578
12	2622	71.6	5314	2	AAQ13576
13	2620.6	71.6	3699	2	AAV14340
14	2609.4	71.3	4001	6	AAD27066
15	2484.6	67.8	3774	6	AAD27062
16	2474.2	67.6	5277	3	AA88110
17	2464.6	67.3	3908	6	ABT08110
18	2451.8	67.0	3928	8	ABT14478
19	2451.8	67.0	3928	8	AAD50628
20	2451.8	67.0	3928	10	ABO84236
21	2448.8	66.9	3956	2	AAV64258

22	2446.8	66.8	3715	6	AAD27064	AAD27064 Plasmid T
23	2446.8	66.8	3927	6	ABT08166	ABT08166 Recombina
24	2314.8	63.2	4088	2	AAV64255	AAV64255 Plasmid P
25	2305.2	62.9	4102	2	AAV64257	AAV64257 Plasmid P
26	2279.2	62.2	4119	6	AAD40772	AAD40772 Plasmid P
27	2271.2	62.0	4059	12	ADN97131	ADN97131 Mammalian
28	2230.6	60.9	5250	10	AAD49957	AAD49957 T4R (T4ran
29	2224.6	60.7	2958	2	AAZ22250	AAZ22250 Nucleotid
30	2216.2	60.5	4205	4	AAD09088	AAD09088 pEA658 P
31	2216.2	60.5	4205	6	AAK99713	AAK99713 DNA of pl
32	2215.4	60.5	4172	2	AAK19901	AAK19901 Plasmid P
33	2213	60.4	4205	4	AAD09087	AAD09087 pEA657 P
34	2213	60.4	4205	6	AAK99700	AAK99700 DNA of pl
35	2202.6	60.1	4454	4	AA806386	AA806386 Vector pG
36	2172.8	59.3	7102	2	AAK13836	AAK13836 Nucleotid
37	2172.8	59.3	7102	6	ABK13259	ABK13259 Yeast pla
38	2172.8	59.3	7333	2	AAK13837	AAK13837 Nucleotid
39	2172.8	59.3	7333	6	ABK13260	ABK13260 Yeast pla
40	2172.8	59.3	7803	6	ABT08177	ABT08177 Recombina
41	2172.8	59.3	8167	6	ABT08178	ABT08178 Recombina
42	2172.6	59.3	10600	11	ADM68432	ADM68432 Celery CE
43	2172.6	59.3	10600	12	ADP26585	ADP26585 TMV CEL I
44	2172.6	59.3	10624	11	ADM68433	ADM68433 Celery CE
45	2172.6	59.3	10624	12	ADP26586	ADP26586 TMV CEL I

ALIGNMENTS

RESULT 1	
ACA55354	
ID	ACA55354 standard; DNA, 3661 BP.
XX	
AC	ACA55354;
XX	
DT	06-JUN-2003 (first entry)
XX	
DE	Transformation vector piggyBAC related plasmid pYL-Bac.
XX	
KW	PiggyBac; transposon; eukaryotic transformation vector; ds;
KW	transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW	circular.
XX	
OS	Synthetic.
XX	
PN	US2002173634-A1.
XX	
PD	21-NOV-2002.
XX	
PF	30-OCT-2001; 2001US-00001189.
XX	
PR	31-OCT-2000; 2000US-0244677P.
XX	
PR	01-NOV-2000; 2000US-0244984P.
XX	
PA	(FRAS/) FRASER M J.
XX	
PA	(LITX/) LI X.
XX	
PA	(BEAM/) BEAM T.
XX	
PA	(HUA/) HUA-VAN A.
XX	
PI	Fraser MJ, Li X, Beam T, Hua-Van A;
XX	
DR	WPI; 2003-352597/33.
XX	
PT	New DNA molecule in the transposon piggyBac, useful for transferring
XX	genes into host cells or embryos for transforming the cells of embryos
PT	that can be used in making transgenic organisms.
XX	
PS	Example 3, Fig 3(C2), 151pp; English.
XX	
CC	The invention describes a DNA molecule comprising at least 163
CC	consecutive nucleotide base pairs of the 3' terminal nucleotide beginning at
CC	the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC	pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac

XX Sequence 3661 BP; 910 A; 910 C; 957 G; 884 T; 0 U; 0 Other;

Query Match 99.6%; Score 3648.4; DB 10; Length 3661;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 1; Gaps 1;

Matches 3660; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 CTTAAATGTAAGCGTTAAATTTGTTAAATTCGCTAAATTTTGTGTTAAATCGCTC 60
DB 1 CTTAAATGTAAGCGTTAAATTTGTTAAATTCGCTAAATTTTGTGTTAAATCGCTC 60
QY 61 ATTTTAAACCAATAGCGCGAATCGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 120
DB 61 ATTTTAAACCAATAGCGCGAATCGCAAAATCCCTTAAATCAAAAGAAATAGACCGA 119
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGAACAAGTCCACTATTAAGAAGCGTGAATC 180
DB 120 GATAGGTTGAGTGTGTTCCAGTTTGAACAAGTCCACTATTAAGAAGCGTGAATC 179
QY 181 CAACGTCAAAAGGCGAAAAACCGTCTATCAGGCGGATGCGCCACTAGTGAAACCATACC 240
DB 180 CAACGTCAAAAGGCGAAAAACCGTCTATCAGGCGGATGCGCCACTAGTGAAACCATACC 239
QY 241 CTATCAATGTTTGGGGTGAAGGCGCTAAAGCACTAAATCGGAACCTTAAAGGAG 300
DB 240 CTATCAATGTTTGGGGTGAAGGCGCTAAAGCACTAAATCGGAACCTTAAAGGAG 299
QY 301 CCCCCATTAGAGCTTGAACGGGAAAGCGCGCAACTGCGGAAAGAAAGGAGGAA 360
DB 300 CCCCCATTAGAGCTTGAACGGGAAAGCGCGCAACTGCGGAAAGAAAGGAGGAA 359
QY 361 AGGAAAGAGCGGGGCTAGAGGCGCTGCAAGTGAAGCGTCAAGTGCAGTAAACGAC 420
DB 360 AGGAAAGAGCGGGGCTAGAGGCGCTGCAAGTGAAGCGTCAAGTGCAGTAAACGAC 419
QY 421 CACACCGCGCGCTTAATGCGCGCTAAGCGCGCTCCATTGCGCTTAAAGCGTGA 480
DB 420 CACACCGCGCGCTTAATGCGCGCTAAGCGCGCTCCATTGCGCTTAAAGCGTGA 479
QY 481 CAACCTTTGGGAAAGGCGGATCGGTGCGGCGCTCTTGCCTATTACGCACTGGCGAAAG 540
DB 480 CAACCTTTGGGAAAGGCGGATCGGTGCGGCGCTCTTGCCTATTACGCACTGGCGAAAG 539
QY 541 GGGATGTGCTGCAAGGCGATTAAAGTGGTAAAGCGCGGTTTCCAGTCAAGAGCTTG 600
DB 540 GGGATGTGCTGCAAGGCGATTAAAGTGGTAAAGCGCGGTTTCCAGTCAAGAGCTTG 599
QY 601 TAAAAAGACGCGCAATGAGCGCGCTGCTCAATCAAGTCTTTTGAACCGCTGAGAGACG 660
DB 600 TAAAAAGACGCGCAATGAGCGCGCTGCTCAATCAAGTCTTTTGAACCGCTGAGAGACG 659
QY 661 GCGAGCTGCGGCTGCAATGTGTTTTCAGCGTGAAGAGCAATGAAGTCTCGACAC 720
DB 660 GCGAGCTGCGGCTGCAATGTGTTTTCAGCGTGAAGAGCAATGAAGTCTCGACAC 719
QY 721 GCTGCAAGACGCGCAATGATTAACCTTAAAGATTAATCATTTGTCAGTCTTAA 780
DB 720 GCTGCAAGACGCGCAATGATTAACCTTAAAGATTAATCATTTGTCAGTCTTAA 779
QY 781 AGATATCATGCGTAAATTTGAACGATGAGATCTGTAAATACGACTCAATAGGCGAAT 840
DB 780 AGATATCATGCGTAAATTTGAACGATGAGATCTGTAAATACGACTCAATAGGCGAAT 839
QY 841 TGGGTACCGGCGCCCTCTGAGAGTGAAGTGAATGATTAATGATTCGATTCCTGC 900
DB 840 TGGGTACCGGCGCCCTCTGAGAGTGAAGTGAATGATTAATGATTCGATTCCTGC 899
```

```
QY 901 AGCCCGGGGATCACTAGTCTAGAGCGCGCCACCGCGGTGAGTCAAGTTTGT 960
DB 900 AGCCCGGGGATCACTAGTCTAGAGCGCGCCACCGCGGTGAGTCAAGTTTGT 959
QY 961 TCCCTTAAAGGCTTAATTAAGTCCGATGCGTCAATTTTACGACATCTTTCTAG 1020
DB 960 TCCCTTAAAGGCTTAATTAAGTCCGATGCGTCAATTTTACGACATCTTTCTAG 1019
QY 1021 GATTATCTAGCTGATCAGATCAATTCGTGCGGCTTTTTCGGCTCAGTATCGCC 1080
DB 1020 GATTATCTAGCTGATCAGATCAATTCGTGCGGCTTTTTCGGCTCAGTATCGCC 1079
QY 1081 CAAGCTGCGCTATCTGGGATTCGGGAGAGAAAGCCCGTTCCTTCCGAGATTG 1140
DB 1080 CAAGCTGCGCTATCTGGGATTCGGGAGAGAAAGCCCGTTCCTTCCGAGATTG 1139
QY 1141 AAGGCGATGGAAGAAGTTTGCGAGATGACTGCTGCAATGAACGTTAGCGGAAAC 1200
DB 1140 AAGGCGATGGAAGAAGTTTGCGAGATGACTGCTGCAATGAACGTTAGCGGAAAC 1199
QY 1201 GCAAGTTTACCATGATGATTCGGGAAAGTGTGCGCATGACGCTTAAACGTTGAAC 1260
DB 1200 GCAAGTTTACCATGATGATTCGGGAAAGTGTGCGCATGACGCTTAAACGTTGAAC 1259
QY 1261 TCGTTCAAGGCGCACTGGGATACCAAGTTCGTGCGGCTTTTCGGAACAAGTTCG 1320
DB 1260 TCGTTCAAGGCGCACTGGGATACCAAGTTCGTGCGGCTTTTCGGAACAAGTTCG 1319
QY 1321 TCAGCCCGAAGCGATCAAGCAACCGGAAACCAACCGGAACTGCGGAGCG 1380
DB 1320 TCAGCCCGAAGCGATCAAGCAACCGGAAACCAACCGGAACTGCGGAGCG 1379
QY 1381 GTGTGCAATTAATGAACGCGGTGCGGCTGAGATTAAGTCAAGCAAGAGCGGATC 1440
DB 1380 GTGTGCAATTAATGAACGCGGTGCGGCTGAGATTAAGTCAAGCAAGAGCGGATC 1439
QY 1441 CTGAGCTGATGCGGCAAAATGACATGATACCCCGTGAAGTTTCCGGGCGGCGCT 1500
DB 1440 CTGAGCTGATGCGGCAAAATGACATGATACCCCGTGAAGTTTCCGGGCGGCGCT 1499
QY 1501 GGGCTAATCATGCTATGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCACA 1560
DB 1500 GGGCTAATCATGCTATGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCACA 1559
QY 1561 CAACATACGAGCGGAAAGATTAAGTGAAGCTCGGGGTGCTTAAGTGAAGTAACT 1620
DB 1560 CAACATACGAGCGGAAAGATTAAGTGAAGCTCGGGGTGCTTAAGTGAAGTAACT 1619
QY 1621 CACATTAATGCTGCTGCGCTCACTGCGCTTTCAGTCAAGTCAAGTCAAGTCAAGT 1680
DB 1620 CACATTAATGCTGCTGCGCTCACTGCGCTTTCAGTCAAGTCAAGTCAAGTCAAGT 1679
QY 1681 GCATTAAATGATGCGGCAACGCGCGGAGAGAGCGGTTTGGTATTTGGGCGCTCTTCCG 1740
DB 1680 GCATTAAATGATGCGGCAACGCGCGGAGAGAGCGGTTTGGTATTTGGGCGCTCTTCCG 1739
QY 1741 TTCTCTGCTCACTGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 1740 TTCTCTGCTCACTGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
QY 1801 CTCAAGGCGGTAATTAAGTTATCAAGATCAAGGGAATGAAGCAAGAAACATGTG 1860
DB 1800 CTCAAGGCGGTAATTAAGTTATCAAGATCAAGGGAATGAAGCAAGAAACATGTG 1859
QY 1861 AGCAAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAG 1920
DB 1860 AGCAAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAG 1919
QY 1921 TAGGCTCGGCGCCCTCTGAGAGTCAAAATTCAGCTCAAGTCAAGGTTGGGAA 1980
DB 1920 TAGGCTCGGCGCCCTCTGAGAGTCAAAATTCAGCTCAAGTCAAGGTTGGGAA 1979
```


Qy 1981 CCGGACAGGACTATTAAGATATACGAGGGCTTCCCGCTGGAGCTCCCTCGTGGCTCTCC 2040
 |||||
 Db 1980 CCGGACAGGACTATTAAGATATACGAGGGCTTCCCGCTGGAGCTCCCTCGTGGCTCTCC 2039
 |||||
 Qy 2041 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGAGACGTCGC 2100
 |||||
 Db 2040 TGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGAGACGTCGC 2099
 |||||
 Qy 2101 GCTTTCATATAGCTACAGCTGTAGTATCTCAGTTCCGCTGTAAGTCGTCCTCAAGCT 2160
 |||||
 Db 2100 GCTTTCATATAGCTACAGCTGTAGTATCTCAGTTCCGCTGTAAGTCGTCCTCAAGCT 2159
 |||||
 Qy 2161 GGGCTGTGTGACGACACCCCGCTTACGCCGACCGGCTGGGCTTATCCGCTAACTATCG 2220
 |||||
 Db 2160 GGGCTGTGTGACGACACCCCGCTTACGCCGACCGGCTGGGCTTATCCGCTAACTATCG 2219
 |||||
 Qy 2221 TCTTAGTCCAAACCCGGTAAAGACAGCTTATCCGCACTGGCAGACGCACTGGTAAACAG 2280
 |||||
 Db 2220 TCTTAGTCCAAACCCGGTAAAGACAGCTTATCCGCACTGGCAGACGCACTGGTAAACAG 2279
 |||||
 Qy 2281 GATTAGCAGACGAGTATGTAGGGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTA 2340
 |||||
 Db 2280 GATTAGCAGACGAGTATGTAGGGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTA 2339
 |||||
 Qy 2341 CCGCTACACTAGAGACAGTATTTGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGG 2400
 |||||
 Db 2340 CCGCTACACTAGAGACAGTATTTGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGG 2399
 |||||
 Qy 2401 AAAAAGATTGGTAGCTCTTGTATCCGCAACAAACACCGCTGTAGCGGTGTTTTT 2460
 |||||
 Db 2400 AAAAAGATTGGTAGCTCTTGTATCCGCAACAAACACCGCTGTAGCGGTGTTTTT 2459
 |||||
 Qy 2461 TGTTCGACGACGAGATTACGCCGACGAAAAAAGATCTCAAGAGATCTTTGATCTT 2520
 |||||
 Db 2460 TGTTCGACGACGAGATTACGCCGACGAAAAAAGATCTCAAGAGATCTTTGATCTT 2519
 |||||
 Qy 2521 TTCTACGGGGCTGACGCTCAGTGGAAAGGAACTCAGCTTAAGGATTTGTGCATGAG 2580
 |||||
 Db 2520 TTCTACGGGGCTGACGCTCAGTGGAAAGGAACTCAGCTTAAGGATTTGTGCATGAG 2579
 |||||
 Qy 2581 ATTATCAAAAAAGATCTTACCTAGATCCTTTAAATTTAAATGAATTTAAATCAAT 2640
 |||||
 Db 2580 ATTATCAAAAAAGATCTTACCTAGATCCTTTAAATTTAAATGAATTTAAATCAAT 2639
 |||||
 Qy 2641 CTAAAGTATATATAGTAACTTGTGCTGACAGTTACCAATGCTTAATCAGTGAAGCAAC 2700
 |||||
 Db 2640 CTAAAGTATATATAGTAACTTGTGCTGACAGTTACCAATGCTTAATCAGTGAAGCAAC 2699
 |||||
 Qy 2701 TATCTCAGCCATCTGTCTATTTGTTTCTATTCATTCATAGTTGCCCTGACTCCCGCTGTAGAT 2760
 |||||
 Db 2700 TATCTCAGCCATCTGTCTATTTGTTTCTATTCATTCATAGTTGCCCTGACTCCCGCTGTAGAT 2759
 |||||
 Qy 2761 AACTACGATACGGAGGGCTTACCATCTGAGCCGACAGTGTCTCAATGATACCGCAGAACCC 2820
 |||||
 Db 2760 AACTACGATACGGAGGGCTTACCATCTGAGCCGACAGTGTCTCAATGATACCGCAGAACCC 2819
 |||||
 Qy 2821 ACCTCACCGGCTCTCAAGTTTATCAGCAATTAACAGCAGCCGAGAGGGCGCAGCCGAG 2880
 |||||
 Db 2820 ACCTCACCGGCTCTCAAGTTTATCAGCAATTAACAGCAGCCGAGAGGGCGCAGCCGAG 2879
 |||||
 Qy 2881 AAGTGTCTTGCACTTTATTCGGCTCTCATTCAGTCTTATTAATTTGTTCCGGGAAGCTAG 2940
 |||||
 Db 2880 AAGTGTCTTGCACTTTATTCGGCTCTCATTCAGTCTTATTAATTTGTTCCGGGAAGCTAG 2939
 |||||
 Qy 2941 AGTAAGTATGTCGCACTTATTAATTTGCGCAAGCTTTGTCGCAATGCTACAGGATCTGT 3000
 |||||
 Db 2940 AGTAAGTATGTCGCACTTATTAATTTGCGCAAGCTTTGTCGCAATGCTACAGGATCTGT 2999
 |||||
 Qy 3001 GGTGTCAAGCTGTGTGTTGTGTATGCTTCAATCAGCTCCGGTTCCCAACGATCAAGCGC 3060
 |||||
 Db 3000 GGTGTCAAGCTGTGTGTTGTGTATGCTTCAATCAGCTCCGGTTCCCAACGATCAAGCGC 3059
 |||||
 Qy 3061 AGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCTCCGATCTGT 3120

Db 3060 AGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCTCCGATCTGT 3119
 |||||
 Qy 3121 TGTTCAGAAAGTAAAGTGTGGCGCAGTGTATTAATCACTCATGTTATGACACACTGCATAATTC 3180
 |||||
 Db 3120 TGTTCAGAAAGTAAAGTGTGGCGCAGTGTATTAATCACTCATGTTATGACACACTGCATAATTC 3179
 |||||
 Qy 3181 TCTTACGTATGATGCAATCCGTAAGATGCTTTCTGTGATCTGTGATGATCACTCAACCAATGC 3240
 |||||
 Db 3180 TCTTACGTATGATGCAATCCGTAAGATGCTTTCTGTGATGATGATCACTCAACCAATGC 3239
 |||||
 Qy 3241 ATTCTGAAATATGATGATGTCGCGCAACGAGTGTGCTTTGCGCGGCTCAATACGGATAA 3300
 |||||
 Db 3240 ATTCTGAAATATGATGATGTCGCGCAACGAGTGTGCTTTGCGCGGCTCAATACGGATAA 3299
 |||||
 Qy 3301 TACCGCGCACTATGACGAACTTTAAAGTCTCATCATTTGAAGAAAGCTTCTTCGGGGCG 3360
 |||||
 Db 3300 TACCGCGCACTATGACGAACTTTAAAGTCTCATCATTTGAAGAAAGCTTCTTCGGGGCG 3359
 |||||
 Qy 3361 AAAAATCTCAAGATCTTACCGCTGTGTGATGATCAATGATGATTAACCACTGTCGACCC 3420
 |||||
 Db 3360 AAAAATCTCAAGATCTTACCGCTGTGTGATGATGATCAATGATGATTAACCACTGTCGACCC 3419
 |||||
 Qy 3421 CAACTGATCTTACGATCTTTTACTTCAACAGGCTTCTGGGTGACCAAAAAACAGGAG 3480
 |||||
 Db 3420 CAACTGATCTTACGATCTTTTACTTCAACAGGCTTCTGGGTGACCAAAAAACAGGAG 3479
 |||||
 Qy 3481 GCAAAATGCGCAAAAAAGGAAATTAAGGGGACACGGAAATGTTGAATCACTCATACTCTT 3540
 |||||
 Db 3480 GCAAAATGCGCAAAAAAGGAAATTAAGGGGACACGGAAATGTTGAATCACTCATACTCTT 3539
 |||||
 Qy 3541 CCTTTTCAATATTTATTAAGCATTTATCAAGGTTATTTGCTCATGAGCGGATACATATT 3600
 |||||
 Db 3540 CCTTTTCAATATTTATTAAGCATTTATCAAGGTTATTTGCTCATGAGCGGATACATATT 3599
 |||||
 Qy 3601 TGAATGATTTAGAAAAATTAACAAATATAGGGGTTCCGCGACATTTCCCGGAAAGTGCC 3660
 |||||
 Db 3600 TGAATGATTTAGAAAAATTAACAAATATAGGGGTTCCGCGACATTTCCCGGAAAGTGCC 3659
 |||||
 Qy 3661 AC 3662
 ||
 Db 3660 AC 3661

RESULT 2
 ADM68463/c
 ID ADM68463 standard; DNA; 3637 BP.
 XX
 AC ADM68463;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Variant green fluorescent protein, GFP, gene construct.
 XX
 KW de; mismatch endonuclease; endonuclease; gene shuffling technology;
 KW single nucleotide polymorphism; cancer susceptibility;
 KW sequence variation redistribution; GFP; green fluorescent protein; gene.
 OS Aequorea victoria.
 OS Synthetic.
 XX
 FN US2003157682-A1.
 XX
 PD 21-AUG-2003.
 XX
 PF 31-JAN-2003; 2003US-00356708.
 XX
 PR 01-FEB-2002; 2002US-0353722P.
 PR 14-MAR-2002; 2002US-00098155.
 PR 01-AUG-2002; 2002US-00211079.
 XX
 PA (PADG/) PADGETT H S.
 PA (VAEW/) VAEMHONGS A A.


```
QY 1620 TCACATTAATTGGCGTGCCTACATGCCCGCTTTCAGTCCGGGAAACCTGTGTCGACG 1679
|||
Db 2043 TCACATTAATTGGCGTGCCTACATGCCCGCTTTCAGTCCGGGAAACCTGTGTCGACG 1984
|||
QY 1680 TGCATTAATGAATCGGCCAACGCGCGGGAGAGAGCGGTTTGCGTAATTTGGGCGCTTTCCG 1739
|||
Db 1983 TGCATTAATGAATCGGCCAACGCGCGGGAGAGAGCGGTTTGCGTAATTTGGGCGCTTTCCG 1924
|||
QY 1740 CTTCCTCGCTCAGTCACTCGCTGCGCTCGGTCTGCTCGGCTCGCGGAGCGGTAATCAAGTC 1799
|||
Db 1923 CTTCCTCGCTCAGTCACTCGCTGCGCTCGGTCTGCTCGGCTCGCGGAGCGGTAATCAAGTC 1864
|||
QY 1800 ACTCAAGGCGGTAATTCGGTTATCCAGAAATCAGAGGGATTAACGACGAAAGAAACATGT 1859
|||
Db 1863 ACTCAAGGCGGTAATTCGGTTATCCAGAAATCAGAGGGATTAACGACGAAAGAAACATGT 1804
|||
QY 1860 GAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGTAAAAAAGCGCGTTCGTGCGTTTTC 1919
|||
Db 1803 GAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGTAAAAAAGCGCGTTCGTGCGTTTTC 1744
|||
QY 1920 ATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAA 1979
|||
Db 1743 ATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAA 1684
|||
QY 1980 ACCGACAGGACTATTAAGATACAGAGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2039
|||
Db 1683 ACCGACAGGACTATTAAGATACAGAGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 1624
|||
QY 2040 CTGTTCCGACCCCTCGCTTACCGGAATACCTGTCCGCTTTCCTTCCTTCGGAAGCGTGG 2099
|||
Db 1623 CTGTTCCGACCCCTCGCTTACCGGAATACCTGTCCGCTTTCCTTCCTTCGGAAGCGTGG 1564
|||
QY 2100 CGCTTTCTCATAGCTCAGCGTGAAGTATCTCAAGTTCGGTGAAGTGTTCCTCCAAAGC 2159
|||
Db 1563 CGCTTTCTCATAGCTCAGCGTGAAGTATCTCAAGTTCGGTGAAGTGTTCCTCCAAAGC 1504
|||
QY 2160 TGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCCGACCGCTGCGCTTAATCCGGTAATCATC 2219
|||
Db 1503 TGGGCTGTGTGCAAGAAACCCCGCTTTCAGCCCGACCGCTGCGCTTAATCCGGTAATCATC 1444
|||
QY 2220 GTCTTGAATCAAAACCCCGTAAGACACGACTTATCGCCACTGGACAGCAGCCTGGTAACA 2279
|||
Db 1443 GTCTTGAATCAAAACCCCGTAAGACACGACTTATCGCCACTGGACAGCAGCCTGGTAACA 1384
|||
QY 2280 GGATTTAGCAGAGCAGGTATGAGCGGTGTCTACAGAGTTCTTGAAGTGTGCTTAATCT 2339
|||
Db 1383 GGATTTAGCAGAGCAGGTATGAGCGGTGTCTACAGAGTTCTTGAAGTGTGCTTAATCT 1324
|||
QY 2340 ACGGCTACCTAGAAAGACATATTTGGTATCTGCGCTCTGTGAGCACTTACCTTCG 2399
|||
Db 1323 ACGGCTACCTAGAAAGACATATTTGGTATCTGCGCTCTGTGAGCACTTACCTTCG 1264
|||
QY 2400 GAAAAAGAGTGTAGCTCTTGATCCGGCAAAACCAACCGCTGTGTGCGGTGGTTCCT 2459
|||
Db 1263 GAAAAAGAGTGTAGCTCTTGATCCGGCAAAACCAACCGCTGTGTGCGGTGGTTCCT 1204
|||
QY 2460 TTGTTTGCAGACAGCATTAACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCT 2519
|||
Db 1203 TTGTTTGCAGACAGCATTAACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCT 1144
|||
QY 2520 TTCTTAGGGGTCTGACGCTCAGTGGAACGAAAACTCAGTTAAGGGATTTTGGTCAATGA 2579
|||
Db 1143 TTCTTAGGGGTCTGACGCTCAGTGGAACGAAAACTCAGTTAAGGGATTTTGGTCAATGA 1084
|||
QY 2580 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAGATTTTAAATCA 2639
|||
Db 1083 GATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAGATTTTAAATCA 1024
|||
QY 2640 TCTAAGTATATATAGATTAACCTTGTGTGACAGTTACCAATGCTTAATCAGTAGGAC 2699
|||
Db 1023 TCTAAGTATATATAGATTAACCTTGTGTGACAGTTACCAATGCTTAATCAGTAGGAC 964
|||
QY 2700 CTATCTCAGGATCTGTCTATTTGTTTCATCATAGTTGCTGACTCCCGCTCGTGTAGA 2759
|||

Db 963 CTATCTCAGGATCTGTCTATTTGTTTCATCATAGTTGCTGACTCCCGCTCGTGTAGA 904
|||
QY 2760 TAACTAGATACGGGAGGGCTTAAACATCTGAGCCCAAGTGTCAATATACCGGAGACC 2819
|||
Db 903 TAACTAGATACGGGAGGGCTTAAACATCTGAGCCCAAGTGTCAATATACCGGAGACC 844
|||
QY 2820 CACGCTCAGCGGCTCCAGATTTTATCAGCAATTAACACGCGCAGGAGGAGGCGAGGCA 2879
|||
Db 843 CACGCTCAGCGGCTCCAGATTTTATCAGCAATTAACACGCGCAGGAGGAGGCGAGGCA 784
|||
QY 2880 GAAATGATCCGCAACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGCGGAAGCTA 2939
|||
Db 783 GAAATGATCCGCAACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGCGGAAGCTA 724
|||
QY 2940 GAGTAAATGATTCGCGCAGTTAATATGTTTGGCAACGTTGTCACATTCCTACAGCATCG 2999
|||
Db 723 GAGTAAATGATTCGCGCAGTTAATATGTTTGGCAACGTTGTCACATTCCTACAGCATCG 664
|||
QY 3000 TGGTGTCAAGCTCGTGGTTGGTATGCGCTTCAATCACTCGGTTCCCAAGATCAAGGC 3059
|||
Db 663 TGGTGTCAAGCTCGTGGTTGGTATGCGCTTCAATCACTCGGTTCCCAAGATCAAGGC 604
|||
QY 3060 GAGTTACATGATCCCGCATGTTGGCAAAAAGGAGTTAGCTCTGCGTCTCGCATCG 3119
|||
Db 603 GAGTTACATGATCCCGCATGTTGGCAAAAAGGAGTTAGCTCTTGGTCTCGCATCG 544
|||
QY 3120 TTGTCAAGATTAAGTTGGCGCAGTGTATCACTCATGATGTATAGCAGCATCTGATATTT 3179
|||
Db 543 TTGTCAAGATTAAGTTGGCGCAGTGTATCACTCATGATGTATAGCAGCATCTGATATTT 484
|||
QY 3180 CTCTTACTGTCAATGCCATCCGTAAGATGCTTTTGTGACTGTGAGTACTCAACCAAT 3239
|||
Db 483 CTCTTACTGTCAATGCCATCCGTAAGATGCTTTTGTGACTGTGAGTACTCAACCAAT 424
|||
QY 3240 CATTCTGAGATATGATATGCGGGACGAGTGTCTTGCAGCGGCTCAATACGAGATA 3299
|||
Db 423 CATTCTGAGATATGATATGCGGGACGAGTGTCTTGCAGCGGCTCAATACGAGATA 364
|||
QY 3300 ATACCGCGCAGCATAGCAGAACTTTAAAGTGTCAATCATTTGAAACGTTCTTCGCGGC 3359
|||
Db 363 ATACCGCGCAGCATAGCAGAACTTTAAAGTGTCAATCATTTGAAACGTTCTTCGCGGC 304
|||
QY 3360 GAAAACTCTAAGGATTTTACCGCTGTGTGAGTCAAGTGTGAATCCCACTGTGCAC 3419
|||
Db 303 GAAAACTCTAAGGATTTTACCGCTGTGTGAGTCAAGTGTGAATCCCACTGTGCAC 244
|||
QY 3420 CCAACGATCTTCAGCATCTTTTAACTTCAACGAGCTTCTGCGGTGAGCAAAAACAGAA 3479
|||
Db 243 CCAACGATCTTCAGCATCTTTTAACTTCAACGAGCTTCTGCGGTGAGCAAAAACAGAA 184
|||
QY 3480 GGCAAAAATGCGCAAAAAAGGAATAAGGCGCACACGGAATGTTGAATCTACTACTCT 3539
|||
Db 183 GGCAAAAATGCGCAAAAAAGGAATAAGGCGCACACGGAATGTTGAATCTACTACTCT 124
|||
QY 3540 TCCTTTTCAATATTAATGAAGCAATTAATGAGGTTATTTGTCTCATGAGCGGATACATAT 3599
|||
Db 123 TCCTTTTCAATATTAATGAAGCAATTAATGAGGTTATTTGTCTCATGAGCGGATACATAT 64
|||
QY 3600 TTGAATGATTTTGAATAATAAATAAATAGGGGTTCCGCGCAATTTCCCGGAAAAATGCG 3659
|||
Db 63 TTGAATGATTTTGAATAATAAATAAATAGGGGTTCCGCGCAATTTCCCGGAAAAATGCG 4
|||
QY 3660 CAC 3662
|||
Db 3 CAC 1
|||

RESULT 3
ADP26616/c
ID ADP26616 standard; DNA; 3637 BP.
XX
AC ADP26616;
```



```
Db 2382 TGGCCATGGAACAGGTAGTTTCCAGTAGTGCAAAATTAATTTAAAGTGAAGTTTCCGTA 2322
Qy 1330 GTCAAGCCGGAAGCGCATAGCAACCCGAACAATACCGCGACAGCCGGAACCTGCCGTGC 1379
Db 2332 TGATGACATACCTTCAACCTCTCCACTGACAGAAAATTTGTGCCATTAACTACCAATC 2263
Qy 1380 GGTGTCAGATTAAAGACAGCGGTGCGCGCTGGGATATTACGTACAGAGACGAGGTAT 1439
Db 2262 TAAATTCACAAGAAAT-----TGGGACAACTTCAGTGAAGAAAGTTC 2224
Qy 1440 CCTGGCTGGATGCGCGCAAAAATGACATGATATACCCCGTAGATTACCCGGCGGCGCCT 1499
Db 2223 TTCTCCTTACTCATCGGATACCCAGCTTTTGTCCCTTATAGTAGAGGTTAATGCGCCT 2164
Qy 1500 TGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGTCACAATTCAC 1559
Db 2163 TGGCGTAATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGTCACAATTCAC 2104
Qy 1560 ACAACATCGAGCCGGAAGCATTAAGTGTAAAGCTGGGGTGCCTTAATGATGAGCTAAC 1619
Db 2103 ACAACATCGAGCCGGAAGCATTAAGTGTAAAGCTGGGGTGCCTTAATGATGAGCTAAC 2044
Qy 1620 TCACATTAAATGCGTTCGCTCACTGCGCGCTTTCCAGTCGCGGAAAACCTGTGTCAGC 1679
Db 2043 TCACATTAAATGCGTTCGCTCACTGCGCGCTTTCCAGTCGCGGAAAACCTGTGTCAGC 1984
Qy 1680 TGCATTAAATGAAATGCGCAACGCGCGGGAAGAGCGGTTTGCTAATGAGCGCTCTTCG 1739
Db 1983 TGCATTAAATGAAATGCGCAACGCGCGGGAAGAGCGGTTTGCTAATGAGCGCTCTTCG 1924
Qy 1740 CTTCCTCGCTCACTGATCTGCTGCGCTGCGCTGCTGCTGCGCTGCGGAGACGCTATCAGTC 1799
Db 1923 CTTCCTCGCTCACTGATCTGCTGCGCTGCGCTGCTGCTGCGCTGCGGAGACGCTATCAGTC 1864
Qy 1800 ACTCAAAAGCGGTAATACGTTTATCAGAGATCAGGGGATTAACAGAGAAAGAAACATGT 1859
Db 1863 ACTCAAAAGCGGTAATACGTTTATCAGAGATCAGGGGATTAACAGAGAAAGAAACATGT 1804
Qy 1860 GAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAGAGCGCGCTGTGCGCTTTTCC 1919
Db 1803 GAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAGAGCGCGCTGTGCGCTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCCCTTGAAGCATCACAAAAATCGACGCTCAAGTCAAGGTGAGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCTTGAAGCATCACAAAAATCGACGCTCAAGTCAAGGTGAGCGAA 1684
Qy 1980 ACCGACAGAGACTTAAGATACAGAGGCTTTCCCGGAAAGCTCCCTGCGGCTCTC 2039
Db 1683 ACCGACAGAGACTTAAGATACAGAGGCTTTCCCGGAAAGCTCCCTGCGGCTCTC 1624
Qy 2040 CTGTTCCGACCCCTGACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCCCTGACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGG 1564
Qy 2100 CGCTTTTCATAGCTCAGGCTGTAGGTATCTCAAGTTCGCTGAGTCTGCTTCCAAAGC 2159
Db 1563 CGCTTTTCATAGCTCAGGCTGTAGGTATCTCAAGTTCGCTGAGTCTGCTTCCAAAGC 1504
Qy 2160 TGGGCTGTGTCAGAAACCCCGCTTACGCGGACCGGCTGCGCTTATCGGTAATATC 2219
Db 1503 TGGGCTGTGTCAGAAACCCCGCTTACGCGGACCGGCTGCGCTTATCGGTAATATC 1444
Qy 2220 GTCTTGAATCAACCCCGGTAAGACAGACTTATCGCACTGACAGACCACTGGTAACA 2279
Db 1443 GTCTTGAATCAACCCCGGTAAGACAGACTTATCGCACTGACAGACCACTGGTAACA 1384
Qy 2280 GGATTACAGAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAAGGTGTGCGCTAATC 2339
Db 1383 GGATTACAGAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAAGGTGTGCGCTAATC 1324
Qy 2340 ACGGCTACACTAGAGAGCACTATTTGGTATCTGCGCTCTGCTGAGCACTTACTTCCG 2399
Db 1323 ACGGCTACACTAGAGAGCACTATTTGGTATCTGCGCTCTGCTGAGCACTTACTTCCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTGAATCCGCAAAACAAACCAACCGCTGTGAGCGGTGTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGAATCCGCAAAACAAACCAACCGCTGTGAGCGGTGTTT 1204
Qy 2460 TTGTTTCAAGCAGACAGATTTCGCGGCAAAAAAAGATCTCAAGAAATCCCTTGAATCT 2519
Db 1203 TTGTTTCAAGCAGACAGATTTCGCGGCAAAAAAAGATCTCAAGAAATCCCTTGAATCT 1144
Qy 2520 TTTCTACGGGCTCAGCGCTCAGTGGAACGAAACTCAGCTTAAGGATTTTGGTCAATGA 2579
Db 1143 TTTCTACGGGCTCAGCGCTCAGTGGAACGAAACTCAGCTTAAGGATTTTGGTCAATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAA 1024
Qy 2640 TCTAAGATATTAAGATTAATAAATTGTCGACATTAACCAATGCTTAATCAAGTACAGC 2699
Db 1023 TCTAAGATATTAAGATTAATAAATTGTCGACATTAACCAATGCTTAATCAAGTACAGC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTTGGTTCATTCATAGTTGCTGACTCCCGCTCGTGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTGGTTCATTCATAGTTGCTGACTCCCGCTCGTGA 904
Qy 2760 TAACTACGATACGGAGAGGCTTACCATCTGCGCCCACTGCTGCAATGATACCGGAGAC 2819
Db 903 TAACTACGATACGGAGAGGCTTACCATCTGCGCCCACTGCTGCAATGATACCGGAGAC 844
Qy 2820 CACGCTACCGGCTCCAGATTTATACCAATTAACCAAGCCAGCCGGAAGGCGCAAGCGCA 2879
Db 843 CACGCTACCGGCTCCAGATTTATACCAATTAACCAAGCCAGCCGGAAGGCGCAAGCGCA 784
Qy 2880 GAAATGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAACCTA 2939
Db 783 GAAATGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGCGGAACCTA 724
Qy 2940 GAGTAAGTATGTCGCAAGTTAATATGTTGCGCAACGTTGTCATGCTACAGGCAATCG 2999
Db 723 GAGTAAGTATGTCGCAAGTTAATATGTTGCGCAACGTTGTCATGCTACAGGCAATCG 664
Qy 3000 TGTGTCAACGCTGCTGTTTGTGATGCTTATCCAGTCTCAGTCTATTAATTTGTCGCGGA 3059
Db 663 TGTGTCAACGCTGCTGTTTGTGATGCTTATCCAGTCTCAGTCTATTAATTTGTCGCGGA 604
Qy 3060 GAGTTACATGATCCCGCATGTTGTCGCAAAAAGCGGTTAGCTCTTCCGCTCCGATCG 3119
Db 603 GAGTTACATGATCCCGCATGTTGTCGCAAAAAGCGGTTAGCTCTTCCGCTCCGATCG 544
Qy 3120 TTGTCAAGATTAAGTTGCGCGCAGTGTATCACTCATGTTATGCGACACTGCAATATT 3179
Db 543 TTGTCAAGATTAAGTTGCGCGCAGTGTATCACTCATGTTATGCGACACTGCAATATT 484
Qy 3180 CTCTTACTGTCAATGCAATCCGTAAGAGCTTTTCTGTGACTGTGTGAGTATCTCAACCAAGT 3229
Db 483 CTCTTACTGTCAATGCAATCCGTAAGAGCTTTTCTGTGACTGTGTGAGTATCTCAACCAAGT 424
Qy 3240 CATTCTGAGATAGTATGAGCGGCAACGAGTTGCTCTTCCCGCGCGCTCAATACGGGATA 3299
Db 423 CATTCTGAGATAGTATGAGCGGCAACGAGTTGCTCTTCCCGCGCGCTCAATACGGGATA 364
Qy 3300 ATACCGGCAACATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTCCGCGGC 3359
Db 363 ATACCGGCAACATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTCCGCGGC 304
Qy 3360 GAAAACTCTAAGATCTTACCGCTGTGTGATTCAGATTGAGATGAACCACTGTGACAC 3419
Db 303 GAAAACTCTAAGATCTTACCGCTGTGTGATTCAGATTGAGATGAACCACTGTGACAC 244
Qy 3420 CCAACTGATCTTCAGCATCTTTTACTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAA 184
```

	Best Local Similarity	83.8%	Pred. No. 0;	Matches 3070;	Conservative	0;	Mismatches 567;	Indels 25;	Gaps 4
Y	1	CTAATGTGAGGCTAATATTTTGTAAATATGCGTAAATTTTGTAAATCAGCTC	60						
b	3637	CTAATGTGAGGCTTAATTTTGTAAATATGCGTAAATTTTGTAAATCAGCTC	3578						
Y	61	ATTTTAAACCATAGGCCGAATAGCGCAAAATCCCTTAATATCAAAGAAATGACCGA	120						
Y	3577	ATTTTAAACCATAGGCCGAATAGCGCAAAATCCCTTAATATCAAAGAAATGACCGA	3518						
b	121	GATAGGGTGAAGTGTTCGAGTTTGAAACAAAGATCCATATTAAGAAAGCTGATC	180						
Y	3517	GATAGGGTGAAGTGTTCGAGTTTGAAACAAAGATCCATATTAAGAAAGCTGATC	180						

[illegible]

Oy 661 GCAGACTGCGGTGCAAATGTGTTTACAGCGCTATGAGCAGATGAAGAATGTCGACAC 720
| | | |
Db 2977 CCGCGGTGGCG-----CCGCTTAAGAACTAGTGATCCCC--CGG 2935
Oy 721 GCTGCATAAATACCACCACTCATATTG

Db	2938		GTCTCAGAGAAATTTCTATTGTTGTAATGATTCATGCGATCGAGACAGATTA	28797
Qy	781		AGATAATCATGCGTAAATATGACGATATGGCATCTGTAAATCGACTCATATATGGCGCAT	840
Db	2878		CAAACTCAGAAAGAACCATGTGTGCAGCGCTTTTCGTTGGGATCTTCCAAAGGCGAGATT	2819
Qy	841		TGGGATCCGGGCCCCCTCGAGGTCGACGGTATCATTAAGCTTAATCGAATTCCTGCG	900
Db	2818		GTGTGCACAGTAATGCTTGTCTGTATTAAGGACAGGGCCATCGCCAAATGGAAGATTTT	2759
Qy	901		AGCCCGGGGAGATCCATAGTTCTAGACGGCGCCACCGCGGTGGAGCTCAGCGTTTGT	960
Db	2758		GTTGATTAATGTGTCTGCTAGTTGAACGAGATTCATTTCAATGTGTGGGGAATTTTGAAGT	2659

2698 TAGCTTGATTCATCTTTTGTGTCGCCGATGATATACATTGTGTGAG-TTATAG 2640
 Db

1021 GGTTATCTAGCTGCATCAGGATCATCTCTGGCCCTCTTTTATATGTTGTTAG- TT TTT

Db 2639 TTGATCTGAGTTGTGTCCGGAATGTTTCATCTCTTTAAATCAATACCTTTTAAAC 2580
Qy 1081 CAAGTCGCGCTATCTGTGGCAATCGGGAGAAAGAAAGCCGTGCTTTTCCGCGAGTTG 1140
Db 2579 TCGATACATTAACAAAGGGATACCTTCAAACTTGACTTCAGCAGCGCTCTTTAGATTTC 2520
Qy 1141 AAGCGGACATGAAAGATTGGCCGAGATGACTGTGTGCAATGACGTTAGAGCGAANAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATATGTGCTTCTGTATATACCTTCCGGCAGTGGCACTTTG 2460
Qy 1201 GCACGTTTACATGATGATTCGGGAAGGTGTGGCCATGACAGCCCTTTAATCGGTGAATGT 1260
Db 2459 AAAAAGTCATGCCGTTTATATATGATCCGGATTAACGGGAAAAGCATTTGAACACANTAAAG 2400
Qy 1261 TCGTTCAAGCCACCTGGGATACCAAGTTGTGCGGCTTTTCCGACACAGTTCCGAGTGG 1320
Db 2399 AAAGTAGTGAACAAGTGTGGCCATGGAAACAGGTATTTTCCAGTGTGCAATAATTA 2340
Qy 1321 TCAGCCGGAAGGCAATCAGCAACCGGAACATACCGGGGACAGCCGGAACGTGCGTGC 1380
Db 2339 AGGTTAGCTTTCCGTTATGTAGCATCACCTTTCACCTTCCATGACAGAAATTTGTC 2280
Qy 1381 GTGTGCAATTAATGACACGGGTGCGCGCTGGGATTTAAGTCAGCAGAGACGGGTATC 1440
Db 2279 CCATTAACATCAACCATCTAATTTCAACAAGAAATGGGACAATCCAGTGAATAAGTTCTT 2220
Qy 1441 CTGGCTGAGTCCCGAGAAATGACATGAAATCCCGTAGTTACCCGCGCGCGCTT 1500
Db 2219 C---CTTTACTCATGCTGATCCAGCTTTTGTTCCTTTAGTAGAGGTTAATATGCGCGCTT 2163
Qy 1501 GGGGTAATCATGTCATGCTGTTTCCGTGTGAAATGTTTATCCGCTCACAAATTTCCA 1560
Db 2162 GGGGTAATCATGTCATGCTGTTTCCGTGTGAAATGTTTATCCGCTCACAAATTTCCA 2103
Qy 1561 CAACATACGACCGGAAACATTAAGTGTAAAGCTGGGTGCTTATGATGAGTAACT 1620
Db 2102 CAACATACGACCGGAAACATTAAGTGTAAAGCTGGGTGCTTATGATGAGTAACT 2043
Qy 1621 CACATTAATTTGGGTGGCTCACTGCGCTTTTCCAGTCGGGAAACTGTGTGTCAGCT 1680
Db 2042 CACATTAATTTGGGTGGCTCACTGCGCTTTTCCAGTCGGGAAAACGTGTGTGTCAGCT 1983
Qy 1681 GCATTAATGAATCGGCAACCGGGGGGAGAGCGGTTTGGGTATGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCAACCGGGGGGAGAGCGGTTTGGGTATGGGCGCTCTTCCGC 1923
Qy 1741 TTCTCTGCTCACTGACTGCTGCGCTCGGTGTGCTGCGCGAGCGGATACACTCA 1800
Db 1922 TTCTCTGCTCACTGACTGCTGCGCTCGGTGTGCTGCGCGAGCGGATACACTCA 1863
Qy 1801 CTCAAAAGCGGTAATACGGTTATTCACAGAAATCAGGGGATTAACGAGAAAGAAATGTG 1860
Db 1862 CTCAAAAGCGGTAATACGGTTATTCACAGAAATCAGGGGATTAACGAGAAAGAAATGTG 1803
Qy 1861 AGCAAAAGGCGAGAAAGGCGAGAAACCGTAAAGGCGCGCTTGTGGGCTTTTTC 1920
Db 1802 AGCAAAAGGCGAGAAAGGCGAGAAACCGTAAAGGCGCGCTTGTGGGCTTTTTC 1743
Qy 1921 TAGGCTCGGCGCTGACGAGCATCACAAAATATGACGCTCAAGTCAGAGGTGGC 1980
Db 1742 TAGGCTCGGCGCTGACGAGCATCACAAAATATGACGCTCAAGTCAGAGGTGGC 1683
Qy 1981 CCGCAGAGAGATTAAGATACAGAGCGTTTCCCGCTGAAAGCTCCCTGTGCGCTTCC 2040
Db 1682 CCGCAGAGAGATTAAGATACAGAGCGTTTCCCGCTGAAAGCTCCCTGTGCGCTTCC 1623
Qy 2041 TGTTCGAGCGCTGCGCTTACCGGATACGTCGCGCTTCCCTTCCGGAAGAGTGGC 2100
Db 1622 TGTTCGAGCGCTGCGCTTACCGGATACGTCGCGCTTCCCTTCCGGAAGAGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCAAGCTGTAGGTATCTCAAGTTCCGTTGAGGTCTTCCCTCCAACT 2160
Db 1562 GCTTTCTCATAGCTCAAGCTGTAGGTATCTCAAGTTCCGTTGAGGTCTTCCCTCCAACT 1503

Qy 2161 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAATATCG 2220
Db 1502 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAATATCG 1443
Qy 2221 TCTTAGTCCAAACCGGTAAGACACGACTTATCCGCACTGGACAGCACACTGGTAACAG 2280
Db 1442 TCTTAGTCCAAACCGGTAAGACACGACTTATCCGCACTGGACAGCACACTGGTAACAG 1383
Qy 2281 GATTAGCAGAGGATATGTAGCGGTGTACAGAGTTCTTGAAGTGTGGCTTAAC 2340
Db 1382 GATTAGCAGAGGATATGTAGCGGTGTGTACAGAGTTCTTGAAGTGTGGCTTAAC 1323
Qy 2341 CCGCTACACTTAAGAGACATTTTGTGATTTGTGGCTCTGCTGAAGCAGTACTTCTGG 2400
Db 1322 CCGCTACACTTAAGAGACATTTTGTGATTTGTGGCTCTGCTGAAGCAGTACTTCTGG 1263
Qy 2401 AAAAAGATTGATGCTTGAATCCGGCAAAACAACACCGCTGTAGCGGTGTTTTT 2460
Db 1262 AAAAAGATTGATGCTTGAATCCGGCAAAACAACACCGCTGTAGCGGTGTTTTT 1203
Qy 2461 TGTTTGACAGCAGATTAACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCTT 2520
Db 1202 TGTTTGACAGCAGATTAACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCTT 1143
Qy 2521 TTTTACGGGGCTGACCGCTCAGTGGAAAGAAATCTACGTTAAGGATTTTGTATGAG 2580
Db 1142 TTTTACGGGGCTGACCGCTCAGTGGAAAGAAATCTACGTTAAGGATTTTGTATGAG 1083
Qy 2581 ATTATCAAAAAGATCTTCACTGATCTTTTAAATTTAAATGAAGTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGATCTTCACTGATCTTTTAAATTTAAATGAAGTTTAAATCAAT 1023
Qy 2641 CTAAAGTATATATAGTAAACCTTGTGTGACAGTTACCAATGCTTAAATCAAGTGGCAC 2700
Db 1022 CTAAAGTATATATAGTAAACCTTGTGTGACAGTTACCAATGCTTAAATCAAGTGGCAC 963
Qy 2701 TATCTCAGCATCTGTCTATTTGTTTCTATCTCATATGTTGCTGACTCCCGCTGTGTAT 2760
Db 962 TATCTCAGCATCTGTCTATTTGTTTCTATCTCATATGTTGCTGACTCCCGCTGTGTAT 903
Qy 2761 AACTAGATACGGGAGGGCTTAACCATCGGCGCCAGTGTGCATTAACCGAGACC 2820
Db 902 AACTAGATACGGGAGGGCTTAACCATCGGCGCCAGTGTGCATTAACCGAGACC 843
Qy 2821 ACGCTCAACGGCTCAAGATTTATCAGCAATTAACACAGCCAGCCGAAAGGCGAGCGAG 2880
Db 842 ACGCTCAACGGCTCAAGATTTATCAGCAATTAACACAGCCAGCGGAGCGAGCGAG 783
Qy 2881 AAGTGTCTGCAACTTTATCCGCTCAATCCAGTCAATTAATTTGTCGGGAAAGCTAG 2940
Db 782 AAGTGTCTGCAACTTTATCCGCTCAATCCAGTCAATTAATTTGTCGGGAAAGCTAG 723
Qy 2941 AGTAAGATTTGCGCAATTAATGTTGGCAACGTTGTTGCCATTTGCTACAGCATCT 3000
Db 722 AGTAAGATTTGCGCAATTAATGTTGGCAACGTTGTTGCCATTTGCTACAGCATCT 663
Qy 3001 GGTGTCAAGCTCGTGTGGTATGAGCTTCAATTCAGCTCGGTTCCCAAGATCAAGGCG 3060
Db 662 GGTGTCAAGCTCGTGTGGTATGAGCTTCAATTCAGCTCGGTTCCCAAGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGGTCTCCGATCT 3120
Db 602 AGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGGTCTCCGATCT 543
Qy 3121 TGTGAGAAATGATGAGTGGCCGAGTGTATACATCTATGTTATGACACATGCAATATTC 3180
Db 542 TGTGAGAAATGATGAGTGGCCGAGTGTATACATCTATGTTATGACACATGCAATATTC 483
Qy 3181 TCTTACTGTATGCAATCCGTAAGATCTTTCTGTGATCTGTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTATGCAATCCGTAAGATCTTTCTGTGATCTGTGAGTACTCAACCAAGTC 423


```

OY 3341 ATTCTGAATATGATGATGCGCGACCGAGTTGCTCTTCCCGGGGCTCAATACGGGATTA 3300
DB 422 ATTCTGAATATGATGATGCGCGACCGAGTTGCTCTTCCCGGGGCTCAATACGGGATTA 363
OY 3301 TACCGGCGCATAGAGAACTTTAAAGTGCATCATGGAAGAAAGTTCTTCGGGGCG 3360
DB 362 TACCGGCGCATAGAGAACTTTAAAGTGCATCATGGAAGAAAGTTCTTCGGGGCG 303
OY 3361 AAAACTCTCAAGAGATCTTCCCGCTGTGATGATCGAGTTGATGTAACCACTCGTGACC 3420
DB 302 AAAACTCTCAAGAGATCTTCCCGCTGTGATGATCGAGTTGATGTAACCACTCGTGACC 243
OY 3421 CAACGATCTTACGATCTTTACTTCAACGAGGTTTCGGGGAGCAAAAACAGAGAG 3480
DB 242 CAACGATCTTACGATCTTTACTTCAACGAGGTTTCGGGGAGCAAAAACAGAGAG 183
OY 3481 GCAAAATGCGCGCAAAAAGGAAATAGGCGCACACGGAATGTTGAATCTCATCTCTT 3540
DB 182 GCAAAATGCGCGCAAAAAGGAAATAGGCGCACACGGAATGTTGAATCTCATCTCTT 123
OY 3541 CTTTTTCAATATATGAAAGCATTTATCAGGGTTATGTTCTCATAGCGGATACATAT 3600
DB 122 CTTTTTCAATATATGAAAGCATTTATCAGGGTTATGTTCTCATAGCGGATACATAT 63
OY 3601 TGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGCACATTTCCCGAAAGTGCC 3660
DB 62 TGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGCACATTTCCCGAAAGTGCC 3
OY 3661 AC 3662
DB 2 AC 1

```

RESULT 5
ADM68448/c
ADM68448 standard; DNA; 3637 BP.

AC ADM68448;

DT 03-JUN-2004 (first entry)

DE Cycle 3 green fluorescent protein, GFP, DNA construct.

XX da; mismatch endonuclease; endonuclease; gene shuffling technology;

KW single nucleotide polymorphism; cancer susceptibility;

KW green fluorescent protein; cycle 3 GFP;

XX Aequorea victoria.

OS Synthetic.

PN US2003157682-A1.

PD 21-AUG-2003.

PF 31-JAN-2003; 2003US-00356708.

PR 01-FEB-2002; 2002US-0353722P.

PR 14-MAR-2002; 2002US-00098155.

PR 01-AUG-2002; 2002US-00211079.

PA (PADG/) PADGETT H S.

PA (VAEW/) VAERHONGS A A.

PA (SMIT/) SMITH M L.

PA (LIND/) LINDBO J A.

PA (FITZ/) FITZMAURICE W P.

PI Padgett HS, Vaerhongs AA, Vojdani FS, Smith ML, Lindbo JA;

PI Fitzmaurice WP;

XX MPI; 2003-766176/72.

PT Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.

XX Example 9; SEQ ID NO 17; 79pp; English.

PS The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the
CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for
CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC DNA construct.

XX Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;

Query Match 72.8%; Score 2664.8; DB 11; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

OY 1 CTAATTTGAACGCTTATATTTTGTAAATTCGGGTTAAATTTTGTAAATTCAGCTC 60
DB 3637 CTAATTTGAACGCTTATATTTTGTAAATTCGGGTTAAATTTTGTAAATTCAGCTC 3578
OY 61 ATTTTTPAACCAATAGGCGCAAAATCCCTTAAATCAAAAAGATAGACCGA 120
DB 3577 ATTTTTPAACCAATAGGCGCAAAATCCCTTAAATCAAAAAGATAGACCGA 3518
OY 121 GATAGGGTTAGTGTGTTTTCAGTTTGAACAAGAGTCCATTTAAAGAGTGAAGTCTC 180
DB 3517 GATAGGGTTAGTGTGTTTTCAGTTTGAACAAGAGTCCATTTAAAGAGTGAAGTCTC 3458
OY 181 CAACGTCAAAAGGCGAAAAACCTCTATCAGGGCGATGCGCATACGTAACCATCACC 240
DB 3457 CAACGTCAAAAGGCGAAAAACCTCTATCAGGGCGATGCGCATACGTAACCATCACC 3398
OY 241 CTATCAAGTTTGTGGGTCAGAGTGCCTTAAGCACTAAATGGAACCTTAAGGAG 300
DB 3397 CTATCAAGTTTGTGGGTCAGAGTGCCTTAAGCACTAAATGGAACCTTAAGGAG 3338
OY 301 CCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAAGTGGCGAAGAGGAGAAAGAA 360
DB 3337 CCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAAGTGGCGAAGAGGAGAAAGAA 3278
OY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTACCGTGGCGTTAAC 420
DB 3277 AGCGAAAGAGCGGGCGCTAGGGCGCTGCAAGTGTAGCGGTACCGTGGCGTTAAC 3218
OY 421 CACACCCCGCGGCTTATATGCGCGCTACAGGGCGGTCCCATTCGCATTCAGGCTGG 480
DB 3217 CACACCCCGCGGCTTATATGCGCGCTACAGGGCGGTCCCATTCGCATTCAGGCTGG 3158
OY 481 CAACGTTGGGAAGGCGATCGGTGCGGCTTTCGCTATTAACCGCATGCGGCAAGG 540
DB 3157 CAACGTTGGGAAGGCGATCGGTGCGGCTTTCGCTATTAACCGCATGCGGCAAGG 3098
OY 541 GGGATGTCTGAAGGCGATTAATTTGGGTTAACGCCAGGGTTTTCCAGTACGACGTTG 600
DB 3097 GGGATGTCTGAAGGCGATTAATTTGGGTTAACGCCAGGGTTTTCCAGTACGACGTTG 3038
OY 601 TAAAACGACGCGCATAGAGCGCGCTCGTTCATTCACGTTTTTGAACCCGTGAGACGG 660
DB 3037 TAAAACGACGCGCATAGAGCGCGCTCGTTCATTCACGTTTTTGAACCCGTGAGACGG 2978
OY 661 GCAGACTCGCGGTGAATATGTTTAAAGCGGTATGAGAGCAATGAATGCTGCACAC 720

Db	2977	CCGCGGTGCGG-----CCGCTCTAGAACTAGTGAATCCCC--CGG	2939
Qy	721	GCTGCAGAAACGCGACGACTAGATTAAACCTAGAAAAGATATATATTTGACGTACGTTAA	780
Db	2938	GCTGCAGGAATTTCTTATTTGTATAGTTATCATCCATGCCATGTGTATATCCACAGACAGTTA	2879
Qy	781	AGATTAATCATCGTAAATAATTGACCGATGGATCTGTAAATACAGACTCACTATAGGCGAAT	840
Db	2878	CAAACTCAAGAAAGGACCAATGTGTATCCACTTTTCGTTGGGATCTTTTCAAAAGGCGAGATT	2819
Qy	841	TGGGTACGGGCCCCCCTCGAGGTCCAGGTATCGATTAACCTGATATCGAATTTCCGCG	900
Db	2818	GTGTCCAGAGGTAAATGTTGTCTGGTAAAGACACGGGCCATTCGCCAATTTGAGATATTTT	2759
Qy	901	AGCCCGGGGGATCCACTAGTTCTTAGAGCGGCGCCACGGCGATGAGCTCCAGCTTTTGT	960
Db	2758	GTTGATATATGGCTCTGTAGTTGTAACGGATTCATCTTCAATGTGTGGCGAAATTTGAAGT	2699
Qy	961	TCCCTTAGTAGGGTTAATTAGATCCATGCGTCAATTTTACGACATATCTTCTAG	1020
Db	2698	TAGCTTTGATTCATTTCTTTGTTGTCTGCGCGTATGTATATACATGTGTAG--TTATAG	2640
Qy	1021	GGTTAATTTACTGTGATGAGATCATATCTGTGGGTCTTTTTCGGCTCAGTATGCGC	1080
Db	2639	TTGTATCTGAAATTTGTGTGTCGAGAAATGTTTCCATCTTTTAAATCATATCTTTTAAAC	2580
Qy	1081	CAAGCTGCGCTATCTGGGCATTCGGGGAGAGAAAGCCGTGCTTTTCCGCGAGGTG	1140
Db	2579	TCGATACGATTTAACAGGGTATCACCTTCAAACTTGATTCAGACAGCGGTCTTTAGTTTC	2520
Qy	1141	AAGCGGATGAAAGATTTGCCGAGATGACTGTCTGTGATTAAGCTTACGGGAAAC	1200
Db	2519	CCGTATCTTTGAAAGATATATAGTGGTTCGTATACATAACCTTGGGACATGCACTCTG	2460
Qy	1201	GCAGTTTACCATGATGATTTGCGGAAGGTGTGGCATACAGCCTTTAAGGTGAATCTGT	1260
Db	2459	AAAAAGTCATCCGTTTCAATATGATCCGGAATAACGGGAAAGCATTTGAACCATTAAG	2400
Qy	1261	TCGTTACAGGCCACCTGGGATACAGTTGTGTGGCGCTTTTCCGGACACAGTCCGGATGG	1320
Db	2339	AAAGTAGTGAAGAATGTTGGCCATGGAACAGTATGTTTCCAGTATGTGCCAATAATTA	2340
Qy	1321	TCAGCCCGAAGCGCATACGAACCCGAACTAATCCGCGGACAGCGCGAATCGCGTCCG	1380
Db	2339	AGGGTAAGCTTTCGTATGTAGCATCACCTTCACTCTCTCACTGACAGAAATTTGTGTC	2280
Qy	1381	GTGTGCAGATTAATGACAGCGGTGCGGCTGGGATATTACGTACCGAGGACGGGTATC	1440
Db	2279	CCATTTAATCATCAACATCTTAATTCAACAAGAAATTTGGGACAACTCCAGTAAAGTTCTTCT	2220
Qy	1441	CTGGCTGGATGCCGAGAAATGGAACATGGAATACCCCGTAATTCGCCGCGGGCGGCTT	1500
Db	2219	C---CTTTACTCATGCGTACCACAGTTTGTCTCCCTTTAATGAGGGTTAATTTGGCGCTT	2163
Qy	1501	GGCGTAATCATGTGTCATAGCTGTTTCTGTGTGAATTTGATATCCGCTCACAATTTCA	1560
Db	2162	GGCGTAATCATGTGTCATAGCTGTTTCTGTGTGAATTTGATATCCGCTCACAATTTCA	2103
Qy	1561	CAACATACGAGCCCGAAGCATTAATGTGAAGCTTGGGTGCTTAATGATGACTAACT	1620
Db	2102	CAACATACGAGCCCGAAGCATTAATGTGAAGCTTGGGTGCTTAATGATGACTAACT	2043
Qy	1621	CACATTAATTTGGGTGCGCTCATGCGCCCTTTCAGTCCGGGAAACCTGTGTGCCAGCT	1680
Db	2042	CACATTAATTTGGGTGCGCTCATGCGCCCTTTCAGTCCGGGAAACCTGTGTGCCAGCT	1983
Qy	1681	GCATTTAATGAATCCGGCAAACCGCGGGGAGAGCGGTTTGGTATTTGGGCGCTCTTCCGC	1740
Db	1982	GCATTTAATGAATCCGGCAAACCGCGGGGAGAGCGGTTTGGTATTTGGGCGCTCTTCCGC	1923
Qy	1741	TTCTCTGCTCACTGATCTGCTGCGTCTGGTCTGCTTCCGCTGCGGCGAGCGGATACGCTCA	1800

Db	1922	TTCTCGCTCACTGACTGCTGCTGCGCTCGGTGCTGAGCTGCGGCGAGCGGTATCAGCTCA	1863
Qy	1801	CTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACGAAAGAACATGTG	1860
Db	1862	CTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACGAAAGAACATGTG	1803
Qy	1861	ACGAAAAGGCCACGCAAAAAGGCCAGGACCGTAAAAAGCCGCGTTGCTGGCGTTTTCCTCA	1920
Db	1802	ACGAAAAGGCCACGCAAAAAGGCCAGGACCGTAAAAAGCCGCGTTGCTGGCGTTTTCCTCA	1743
Qy	1921	TAGGCTCGGCCCCCGTGACGAGCATCACAAAAATGACGCTCAAGTCACAGGTGGCGAAA	1980
Db	1742	TAGGCTCGGCCCCCGTGACGAGCATCACAAAAATGACGCTCAAGTCACAGGTGGCGAAA	1683
Qy	1981	CCCGACAGGACTATAAGATACACGAGCGTTTCCCGCTGAAAGCTCCCTGTCGCTCTCC	2040
Db	1682	CCCGACAGGACTATAAGATACCAAGCGTTTCCCGCTGAAAGCTCCCTGTCGCTCTCC	1623
Qy	2041	TGTTCCGACCTGCGCGCTTACCGGATACCTGTGTCGCGCTTTCCTTCGGAAGCGTGGC	2100
Db	1622	TGTTCCGACCTGCGCGCTTACCGGATACCTGTGTCGCGCTTTCCTTCGGAAGCGTGGC	1563
Qy	2101	GCTTCTCATACACTACGCTGATAGATCTCAGTTCCGTTGATGTCGTTCCGTCGAAGCT	2160
Db	1562	GCTTCTCATACACTACGCTGATAGATCTCAGTTCCGTTGATGTCGTTCCGTCGAAGCT	1503
Qy	2161	GCGCTGTGTGACAGAACCCCGCTTCAGGCCACGCTGCGGCTTATCCGTTAACTATCG	2220
Db	1502	GCGCTGTGTGACAGAACCCCGCTTCAGGCCACGCTGCGGCTTATCCGTTAACTATCG	1443
Qy	2221	TCTTAGTCCAAACCCGGTAAAGACACGACTTATCGCACCTGGACAGACCACTGTTAACG	2280
Db	1442	TCTTAGTCCAAACCCGGTAAAGACACGACTTATCGCACCTGGACAGACCACTGTTAACG	1383
Qy	2281	GATTTGACAGAGGAGTATGTATGGGGGGGCTACAAAGTCTTGAAGTGGTGGCTTAACTA	2340
Db	1382	GATTTGACAGAGGAGTATGTATGGGGGGGCTACAAAGTCTTGAAGTGGTGGCTTAACTA	1323
Qy	2341	CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCG	2400
Db	1322	CGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCG	1263
Qy	2401	AAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGATGGCGTGGTTTTT	2460
Db	1262	AAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGATGGCGTGGTTTTT	1203
Qy	2461	TGTTTGACAGCAGCATTTACGCGCAGAAAAAAGATCTCAAGAAAGATCCTTTGATCTT	2520
Db	1202	TGTTTGACAGCAGCATTTACGCGCAGAAAAAAGATCTCAAGAAAGATCCTTTGATCTT	1143
Qy	2521	TTCTACGGGGTGTGACGCTCAGGTGGAAGAAAACCTCAGTTAAGGATTTTGGTACATGAG	2580
Db	1142	TTCTACGGGGTGTGACGCTCAGGTGGAAGAAAACCTCAGTTAAGGATTTTGGTACATGAG	1083
Qy	2581	ATTATCAAAAAGGATCTTCACTAGATCCTTTAAATTTAAAAATGAAAGTTTTAAATCAAT	2640
Db	1082	ATTATCAAAAAGGATCTTCACTAGATCCTTTAAATTTAAAAATGAAAGTTTTAAATCAAT	1023
Qy	2641	CTAAAGTATATATGATTAACCTTGGTCTGACAGTTAACCAATGCTTAAATCAGTAGGACCC	2700
Db	1022	CTAAAGTATATATGATTAACCTTGGTCTGACAGTTAACCAATGCTTAAATCAGTAGGACCC	963
Qy	2701	TATCTCAGCGATCTGTCTAATTCGTTCAATCCATAGTGTGATGCTGATGCTGATGAT	2760
Db	962	TATCTCAGCGATCTGTCTAATTCGTTCAATCCATAGTGTGATGCTGATGCTGATGAT	903
Qy	2761	AACTACGATACGGGAGGCGCTTACCATCTGGCCCCAGTGTGAAATGATACCGCGAGACCC	2820
Db	902	AACTACGATACGGGAGGCGCTTACCATCTGGCCCCAGTGTGAAATGATACCGCGAGACCC	843
Qy	2821	ACGCTCACCGGCTCCAGATTTATCAGCAATTAACCGACCGACCGGAGGGCCGAGCGCAG	2880
Db	842	ACGCTCACCGGCTCCAGATTTATCAGCAATTAACCGACCGGAGGGCCGAGCGCAG	783

QY	2881	AAAGTGTCTGCAACTTATATCCGCCCTCCATCCATCTATTAATTGTTGCGGGAAAGCTAG	2941
Db	782	AAAGGTCCTCGCAACTTATATCCGCCCTCCATCCATCTATTAATTGTTGCGGGAAAGCTAG	723
QY	2941	AGTAAGTAAGTGGCCAGTTAATAATGTTTGGCGCAAGTGTGGCAATGTCACAGCATGT	3000
Db	722	AGTAAGTAAGTGGCCAGTTAATAATGTTTGGCGCAAGTGTGGCAATGTCACAGCATGT	663
QY	3001	GGTGTACGCTCGTGGTGGTATGGGCTTAATCAGCTCCGGTTCACACATCAAGGCG	3060
Db	662	GGTGTACGCTCGTGGTGGTATGGGCTTAATCAGCTCCGGTTCACACATCAAGGCG	603
QY	3061	AGTTACATGATCCCCCAGATGTTGTGCACAAAAAGCGGTTAAGCTCCTGGGTCCGATCGT	3120
Db	602	AGTTACATGATCCCCCAGATGTTGTGTGCACAAAAAGCGGTTAAGCTCCTGGGTCCGATCGT	543
QY	3121	TGTCAAGTAAGTTGGCCGACAGTGTATACATAGTGTATAGCAGACACTGCATAATTC	3180
Db	542	TGTCAAGTAAGTTGGCCGACAGTGTATACATAGTGTATAGCAGACACTGCATAATTC	483
QY	3181	TCTTACTGTATGCAATCCGCTTAAGATGCTTTTGTGTGACTGTAGTAACTCAACAAAGTC	3240
Db	482	TCTTACTGTATGCAATCCGCTTAAGATGCTTTTGTGTGACTGTAGTAACTCAACAAAGTC	423
QY	3241	ATTCTGGAATAGTGTATGCGGCGACCGAGTTGCTTTGCCCGGCTCAATACGGGATTA	3300
Db	422	ATTCTGGAATAGTGTATGCGGCGACCGAGTTGCTTTGCCCGGCTCAATACGGGATTA	363
QY	3301	TACCGGCGCACATAGAGAACTTTAAAGTGTCTATCATTTGAAAAAGTTCTTGCGGGCG	3360
Db	362	TACCGGCGCACATAGAGAACTTTAAAGTGTCTATCATTTGAAAAAGTTCTTGCGGGCG	303
QY	3361	AAAACCTCTAAGAGTCTTACCGCTGTGTGAGATCCAGTCAAGTGAATACCACTGTGCACC	3420
Db	302	AAAACCTCTAAGAGTCTTACCGCTGTGTGAGATCCAGTCAAGTGAATACCACTGTGCACC	243
QY	3421	CAACTGATCTTCAGATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAAAACGAAG	3480
Db	242	CAACTGATCTTCAGATCTTTTACTTTTACACAGCGTTTCTGGGTGAGCAAAAAACGAAG	183
QY	3481	GGAATAATCCCGCAAAAAAGGAATAAGGCGCACAGGAAATGTTGAATACTCATCTCTT	3540
Db	182	GGAATAATCCCGCAAAAAAGGAATAAGGCGCACAGGAAATGTTGAATACTCATCTCTT	123
QY	3541	CCTTTTCATATTAATTAAGAACTTATATAGGGTATATGTCTCATAGCGGATACATATT	3600
Db	122	CCTTTTCATATTAATTAAGAACTTATATAGGGTATATGTCTCATAGCGGATACATATT	63
QY	3601	TGAATGATTTAATAAAATPAACAATATAGGGGTTCCGCGACATTTTCCCGAAAGTCC	3660
Db	62	TGAATGATTTAATAAAATPAACAATATAGGGGTTCCGCGACATTTTCCCGAAAGTCC	3
QY	3661	AC 3662	
Db	2	AC 1	

Result	6
ADP26601/C	
ID	ADP26601 standard; DNA; 3637 BP.
XX	
AC	ADP26601;
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	Green fluorescent protein (GFP) plasmid DNA #2.
XX	
KW	Sequence variation; heteroduplex; transcription; DNA integration;
KM	ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX	
OS	Aequorea victoria.
OS	Synthetic.

```

XX      US2004110130-A1.
PN      10-JUN-2004.
XX      PD
XX      XX
XX      25-OCT-2002; 2002US-00280913.
PF      XX
XX      02-FEB-2001; 2001US-026386P.
PR      14-FEB-2001; 2001US-026878P.
PR      01-FEB-2002; 2002US-0006875P.
PR      08-AUG-2002; 2002US-0402342P.
XX      XX
XX      (LARG-) LARGE SCALE BIOLOGY CORP.
PA      XX
PI      Padgett HS, Lindbo JA, Filtzmaurice WP;
XX      WP1; 2004-440326/41.
XX      DR
XX      XX
XX      PT Redistributing sequence variations between non-identical polynucleotide
PT      sequences, useful for generating improved polynucleotide having a desired
XX      characteristic, comprises making a heteroduplex and introducing a nick.
XX      Example 5; SEQ ID NO 17; 75pp; English.
XX      XX

```

The invention relates to an in vitro method of redistributing sequence variations between non-identical polynucleotide sequences, comprising making a heteroduplex polynucleotide from two non-identical polynucleotides, introducing a nick in the second strand at or near a base pair mismatch site, removing the mismatched base(s) from the mismatch site where the nick occurred and using the first strand as a template to replace the removed base(s) with bases that complement the base(s) in the first strand. The invention also relates to an in vitro method of making a population of sequence variants that complement the polynucleotide sequence, obtaining a polynucleotide sequence encoding a desired functional property and identifying a reassorted DNA molecule encoding a protein with a desired functional property. The method is useful for generating an improved polynucleotide sequence or a population of improved polynucleotide sequences possessing at least one desired phenotypic characteristic (e.g., promotes transcription of linked polynucleotides), where such polynucleotides are useful for expression from a plant, animal, fungal, yeast, or bacterial expression vector, for integration to form a transgenic plant, animal or microorganism, and for expression of a ribozyme. This sequence represents DNA used in the scope of the invention.

sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other

Query Match	72.8%;	Score 2664.8;	DB 12;	Length 3637;
Best Local Similarity	83.8%;	Pred. No. 0;		
Matches 3070;	Conservative			

conservative 0; Mismatches 567; Indels 25; Gaps 4;

Db
3637
CGTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60

555 C A A A T G T A A G C G T T A A T T T T G T T A A A T T C G C G T T A A T T T T G T T A A T C A G C T C 3578

61 AATTTTAAACCAATAGCCGAAATCGCAAAATCCCTATTAATCAAAAGATAGACCGA 120

3577 ATTTTTCACCATAGCGGAATCGGCAAAATCCCTTATAATCAAAAGATAGACCGA 3518

121 GATAGGGTTGAGTCTTCTCCAGTTTGGACAAGACTCCATATTAAAGAACTGGACTCTC 100

Db 3517 GATAGGTTGACTGTTGTTCCAGTTTGGACAACAGCTCCACTATTAAAGAACCTCCCTCTG 180

Qy 181 CAAGTCAAGGGCGAAAAACCGTCATCAGGGGCAATGGCCCACTTGCGAATCGGACIC 3458

[illegible]

QY 241 CTATCAATTCTTGGGTCAGCGCCGAAACCTCACC 3398
.....TTCGCCGAAGCCACACTACGTGTAACCATCACC

3397 CTAACTAAGCTTTTGGGCTGACCTGAGGAG 300

301 CCCCCGATTTCAGACGmnmA CCCCCA
3338

.....TAAAGCCTGACGGGAAAGCCCGGCAACGTTGGCGAGAAAGCAAGGGACAA 360

Dh 3337 CCCCCGATTTAGAGCTTGACGCGGGAAAGCCGGCAACGTGGCGAAGAAAGAAAGGAAAGA 3278
Qy 361 AGCGAAAGAGGCGGCGCTAGGCGCGCTGGCAAGTGTAGCGGTCAAGCTGCGGTGAACAC 420
Db 3277 AGCGAAAGAGGCGGCGCTAGGCGCGCTGGCAAGTGTAGCGGTCAAGCTGCGGTGAACAC 3218
Qy 421 CACACCCGCGGCTTAAATGCGCGCTACAGGGCGCGTCCATTGCGCAATTCAGGCTGG 480
Db 3217 CACACCCGCGGCTTAAATGCGCGCTACAGGGCGCGTCCATTGCGCAATTCAGGCTGG 3158
Qy 481 CAACGTGTGGAAAGGCGCATCGGTGCGGGCTCTTCGCTAATTAACGCAAGTGGCGAAAG 540
Db 3157 CAACGTGTGGAAAGGCGCATCGGTGCGGGCTCTTCGCTAATTAACGCAAGTGGCGAAAG 3098
Qy 541 GGGATGTCTGCAAGCGCATTAAGTTGGGTAAAGCCGAGGGTTTTCCCAAGTCAAGCTTG 600
Db 3097 GGGATGTCTGCAAGCGCATTAAGTTGGGTAAAGCCGAGGGTTTTCCCAAGTCAAGCTTG 3038
Qy 601 TAAACGACGCGCACTGAGCGCGCTCGTTCACTTCAACGTTTTGAACCCGCGTGAAGAGG 660
Db 3037 TAAACGACGCGCACTGAGCGCGCTCGTTAATGCACTATAGGCGCAATTTGAGCTCA 2978
Qy 661 GCAGACTGCGCGTGCMAATGTGTATTGAAGCGGTGATGAGCAGATGAAGATGCTGAGAC 720
Db 2977 CCGCGGTGCGG-----CGGCTAAGACTAGTGGATCCC--CGG 2939
Qy 721 GCTGCAAGAACGCGCACTAGATTAACCTAGAAAGATATCATATTGTGACGTACGTTAA 780
Db 2938 GCTGCAAGAAATCTTAATTGTATAGTTATCATTCATCCATGTCATATCCAGCAGACAGTTA 2879
Qy 781 AGATATCATGCGTAAATTTGACGAGTGGATCTGTAATAGACACTCATATAGGCGCAT 840
Db 2878 CAATCTAGAAAGACATGTGTGACGCTTTTCGTTGGAGTCTTTCGAAAGGCGAGATT 2819
Qy 841 TGGGTACCGGCGCCCTCGAGGTGACGCGTATGATAGCTTGATATCGAATTCCTGTC 900
Db 2818 GTGTGACAGGTATGTGTGTCTGTGTAAGAACAGGGCAATCGCAATTTGAGATTTT 2759
Qy 901 AGCCGGGGGATCCACTAGTCTAAGCGGCGCCACCGCGGTGAGCTCCAGCTTTTGT 960
Db 2758 GTTGAATATGTCTGTCTAGTTGAAACGATTCATCTTCAATGTTGGGAAATTTTGAAGT 2699
Qy 961 TCCCTTAGTGAGGGTAAATTAAGATCCATGCGTCAATTTTACGACATATCTTTAG 1020
Db 2698 TAGCTTATTCATCTTTTGTGTGTGTCTGCGGTATGATATCATTTGTGTAG--TTAATG 2640
Qy 1021 GGTATATCTAGCTGATCAGATCATATCGTGGGCTCTTTTTCGCGCTCAGTCAATCCG 1080
Db 2639 TTGTACTGAGTTGTGTGCGAGATGTTTCCATCTTCTTAAATCAATACCTTTTAAC 2580
Qy 1081 CAAGCTGCGCTATCTGGGCAATCGGGAGAAAGAACCCGTGCTTTTCCGCGAGTTG 1140
Db 2579 TCGATACATTAACAAGGATACACCTTCAAACTTGACTTGAGCAGCGGTCTTGTATGTTG 2520
Qy 1141 AAGCGCATGAAAGATTTGCCGAGATGACTGTGCTGATTAACGTTAGCGAAGAAC 1200
Db 2519 CCGTCATCTTGAAGATATAGTGTCTCTGTATATTAACCTTCCGAGTACGACTTTG 2460
Qy 1201 GCACGTTTACATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTAAACGCTGAATGT 1260
Db 2459 AAAAAGTATGCGTTTATATGATCCGGAATACGGAAAGACATTTGAACACATTAAG 2400
Qy 1261 TCGTTACGCGCACTGGGATACCAATGCTGTGCGGCTTTTCCGACACAGTTCCGAGTGG 1320
Db 2399 AAAGTATGACAAAGTGTGGCCATGGAACAGGTATTTTCCAGTGTGCAATAAATTA 2340
Qy 1321 TCAGCCGAAAGCATACGAACCCGAACATTAACGGGAGACCGCGGAATCGCGTGGC 1380
Db 2339 AGGTATGCTTTCGTAATGATACATCACTTCCATCGACAAATTTTGTGC 2280
Qy 1381 GTGTGCAATTAATGACAGCGGTGCGGCTGGGATATTACGTACGACGAGGCGGATTC 1440
Db 2279 CCATTAACATCACTTAATTCACAAAGAAATTTGGGACAACTCCAGTGAAGATCTTCT 2220

Qy 1441 CTGCTGATGCCCGCAAAATGCAATCGAATACCCCGTAGTACCCGCGGCGCGCTT 1500
Db 2219 C---CTTATCTCATCGATACCCAGCTTTTGTCTCTTATAGAGGTTAATGTGCGCTT 2163
Qy 1501 GCGGTATCATGTGATCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCACAAATTCACA 1560
Db 2162 GCGGTATCATGTGATCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCACAAATTCACA 2103
Qy 1561 CAACATACGACCCGGAAGCATTAAGTGAAGCTGGGCTGCTTAATGATGACTAACT 1620
Db 2102 CAACATACGACCCGGAAGCATTAAGTGAAGCTGGGCTGCTTAATGATGACTAACT 2043
Qy 1621 CACATTAATTTGCGTTGCGCTCACTGCGCTTTTCCAGTGGGAAACCTGTGTGCACT 1680
Db 2042 CACATTAATTTGCGTTGCGCTCACTGCGCTTTTCCAGTGGGAAACCTGTGTGCACT 1983
Qy 1681 GCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATTGGGCGCTCTTCCGC 1923
Qy 1741 TTCTCGCTCACTGACTGCTGCGCTGCTGCTTCCGCTGCGGCGAGCGGTATCAGTCA 1800
Db 1922 TTCTCGCTCACTGACTGCTGCGCTGCTGCTTCCGCTGCGGCGAGCGGTATCAGTCA 1863
Qy 1801 CTCAAGCGGTAATACGTTATTCACAGAAATCAGGGGATTAACGCAAGAAAGAACATGTG 1860
Db 1862 CTCAAGCGGTAATACGTTATTCACAGAAATCAGGGGATTAACGCAAGAAAGAACATGTG 1803
Qy 1861 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAGAAAGCGCGTGTGCTGGCGTTTTC 1920
Db 1802 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAGAAAGCGCGTGTGCTGGCGTTTTC 1743
Qy 1921 TAGCTCGCGCCCTGACGAGCATCAAAATTCAGACCTCAAGTCAAGAGTGGCGAA 1980
Db 1742 TAGCTCGCGCCCTGACGAGCATCAAAATTCAGACCTCAAGTCAAGAGTGGCGAA 1683
Qy 1981 CCGCAACGAGATTAATAAGATTCAGAGGCTTCCCTCGTGAAGCTCCCTCGCTCTCC 2040
Db 1682 CCGCAACGAGATTAATAAGATTCAGAGGCTTCCCTCGTGAAGCTCCCTCGCTCTCC 1623
Qy 2041 TGTTCGACCTGCGCTTACCGGATACGTGTCCGCTTTCCTTCGGAACGATGAC 2100
Db 1622 TGTTCGACCTGCGCTTACCGGATACGTGTCCGCTTTCCTTCGGAACGATGAC 1563
Qy 2101 GCTTTCATATGCTCAAGCTGTAGTATCTCAATTCGATGAGTGTGCTTCCCAACT 2160
Db 1562 GCTTTCATATGCTCAAGCTGTAGTATCTCAATTCGATGAGTGTGCTTCCCAACT 1503
Qy 2161 GGGCTGTGTGACGAACCCCGTTCAAGCCGACCGCTGCGCTTATCCGTAACATACTG 2220
Db 1502 GGGCTGTGTGACGAACCCCGTTCAAGCCGACCGCTGCGCTTATCCGTAACATACTG 1443
Qy 2221 TCTTAGTCCAACCCGGTAAGACAGACTTATGCGCATGCGACAGCACTGTAACAG 2280
Db 1442 TCTTAGTCCAACCCGGTAAGACAGACTTATGCGCATGCGACAGCACTGTAACAG 1383
Qy 2281 GATTAGCAGACGAGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGTGGCTAACTA 2340
Db 1382 GATTAGCAGACGAGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGTGGCTAACTA 1323
Qy 2341 CCGCTACACTAAGAAAGACATTTTGTGATCTGTGCTGTGAAGCAATTAACCTTCCG 2400
Db 1322 CCGCTACACTAAGAAAGACATTTTGTGATCTGTGCTGTGAAGCAATTAACCTTCCG 1263
Qy 2401 AAAAAGTGTGATCTTGAATCGGCAACAAACACCGCTGTGACGCTGTTT 2460
Db 1262 AAAAAGTGTGATCTTGAATCGGCAACAAACACCGCTGTGACGCTGTTT 1203
Qy 2461 TGTTTGACAGCAGATTAACGCGCAGAAAAAAGATTCACAAAGATCTTTGATCTT 2520
Db 1202 TGTTTGACAGCAGATTAACGCGCAGAAAAAAGATTCACAAAGATCTTTGATCTT 1143

QY 2521 TTCTACGGGGCTGACGCTCAGTGGAACGAAATCAGTTAAGGATTTGGTCATGAG 2580
DB 1142 TTCTACGGGGCTGACGCTCAGTGGAACGAAATCAGTTAAGGATTTGGTCATGAG 1083
QY 2581 ATTATCAAAAAGATCTTCACCTGATACCTTTAAATTAATAAGATTTAAATCAAT 2640
DB 1082 ATTATCAAAAAGATCTTCACCTGATACCTTTAAATTAATAAGATTTAAATCAAT 1023
QY 2641 CTTAAGTATATATAGTAAGTAACCTTGCTGACAGTTACCAATGCTTAATACAGTAGGACCC 2700
DB 1022 CTTAAGTATATATAGTAAGTAACCTTGCTGACAGTTACCAATGCTTAATACAGTAGGACCC 963
QY 2701 TATCTACGCGATCTGCTATATTTGCTTATCCATCCATGCTGACCTCCCGCTGATAGAT 2760
DB 962 TATCTACGCGATCTGCTATATTTGCTTATCCATCCATGCTGACCTCCCGCTGATAGAT 903
QY 2761 AACTACGATACGGGAGGGGCTTACATCTGGCCCCAGTGTCTCAATGATACCGGAGACCC 2820
DB 902 AACTACGATACGGGAGGGGCTTACATCTGGCCCCAGTGTCTCAATGATACCGGAGACCC 843
QY 2821 AGCTCACCGGCTCCAGATTTATACGCAATTAACCAACCGGAGAGGGCCGAGCGCAG 2880
DB 842 AGCTCACCGGCTCCAGATTTATACGCAATTAACCAACCGGAGAGGGCCGAGCGCAG 783
QY 2881 AAGTGTCCCTGCAATTTATCCGCTCCATCCATGCTTAATTTGTTGGCGGAGAGCTAG 2940
DB 782 AAGTGTCCCTGCAATTTATCCGCTCCATCCATGCTTAATTTGTTGGCGGAGAGCTAG 723
QY 2941 AGTAAGTATGTCGCAATTAATGTTGCGCAACGTTGTCATTTGCTTACAGGATCGT 3000
DB 722 AGTAAGTATGTTGCGCAATTAATGTTGCGCAACGTTGTCATTTGCTTACAGGATCGT 663
QY 3001 GGTGTACGCTCGCTGCTTGTGTAAGCTTATACAGCTCCGCTCCCAACGATCAAGGCG 3060
DB 662 GGTGTACGCTCGCTGCTTGTGTAAGCTTATACAGCTCCGCTCCCAACGATCAAGGCG 603
QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAGGGGTTAGCTTCCGCTCCGATCGT 3120
DB 602 AGTTACATGATCCCCCATGTTGTGCAAAAAGGGGTTAGCTTCCGCTCCGATCGT 543
QY 3121 TGTGAGAAATGTTGGCGGCACTGTTATCACTCATGTTATGCGACAGCTGCAATATTC 3180
DB 542 TGTGAGAAATGTTGGCGGCACTGTTATCACTCATGTTATGCGACAGCTGCAATATTC 483
QY 3181 TCTTACTGTATGCGCATCCGTAAGATGCTTTCTGTGATCGTGAAGTACTCAACCAAGTC 3240
DB 482 TCTTACTGTATGCGCATCCGTAAGATGCTTTCTGTGATCGTGAAGTACTCAACCAAGTC 423
QY 3241 ATTCTGAGAAATGTTGGCGGCACTGTTATCACTCATGTTATGCGACAGCTGCAATATTC 3300
DB 422 ATTCTGAGAAATGTTGGCGGCACTGTTATCACTCATGTTATGCGACAGCTGCAATATTC 363
QY 3301 TACCGCGGCAATGCGAACTTTTAAAGTGTCTCATCATTTGGAAGAAAGCTTCTCGGGCG 3360
DB 362 TACCGCGGCAATGCGAACTTTTAAAGTGTCTCATCATTTGGAAGAAAGCTTCTCGGGCG 303
QY 3361 AAAAATCTCAAGATCTTACCGGCTGTAAGTCAATGATGTAACCACTCGTGCACC 3420
DB 302 AAAAATCTCAAGATCTTACCGGCTGTAAGTCAATGATGTAACCACTCGTGCACC 243
QY 3421 CAACTGATCTTACGATCTTTTACTTTCACAGCGTTTCTGGGTAGGAGAAACAGGAG 3480
DB 242 CAACTGATCTTACGATCTTTTACTTTCACAGCGTTTCTGGGTAGGAGAAACAGGAG 183
QY 3481 GCAAAATGCGCAAAAAGGGAATTAAGGCGCACAGGAAATGTTGAATACTCATCTT 3540
DB 182 GCAAAATGCGCAAAAAGGGAATTAAGGCGCACAGGAAATGTTGAATACTCATCTT 123
QY 3541 CTTTTTCAATTAATTAAGATTTATCAGGTTATGTTCTCATGAGCGGATACATAT 3600
DB 122 CTTTTTCAATTAATTAAGATTTATCAGGTTATGTTCTCATGAGCGGATACATAT 63
QY 3601 TGAATGTAATTAAGAAAAATAACAAATAGGGGTTCCGCGCAATTTCCCGAAAGTGC 3660

DB 62 TGAATGTAATTAAGAAAAATAACAAATAGGGGTTCCGCGCAATTTCCCGAAAGTGC 3
QY 3661 AC 3662
DB 2 AC 1

RESULT 7
ABX14570/C
ID ABX14570 standard; DNA; 3637 BP.
AC ABX14570;
XX
XX 14-MAR-2003 (first entry)
DT
DE Plasmid pBSMTGFP DNA.
XX
XX
XX Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;
KW population diversity; exonuclease; complementarity; ds.
XX
OS Synthetic.
EN
XX WO200279468-A2.
XX
XX 10-OCT-2002.
PD
PF 01-FEB-2002; 2002WO-US003055.
XX
XX 02-FEB-2001; 2001US-026386P.
PR 14-FEB-2001; 2001US-0268785P.
XX
XX (LARG-) LARGE SCALE BIOLOGY CORP.
PA
XX
PI Padgett HS, Fitzmaurice WP, Lindo JA;
DR WPI; 2003-092898/08.
XX
XX
XX Increasing complementarity by making sequence variants from heteroduplex
PT polynucleotides, useful for producing improved gene products from
PT randomly mutated genes.
XX
XX
XX Example 5; Page 98-99; 105p; English.
XX
XX The invention relates to an in vitro method of making sequence variants
CC from at least one heteroduplex polynucleotide that has at least two-non
CC complementary nucleotide base pairs. The method comprises combining a
CC heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA
CC sufficient time for the percentage of complementarity to increase, and allowing
CC one or more variants are made. The method can also be used to increase
CC diversity in a population of sequences. The method is useful for
CC producing improved gene products from randomly mutated genes or from
CC expression from a suitable plant, animal, fungal, yeast or bacterial
CC recombinants for generating novel polynucleotides from parental templates
CC that are more diverse and with a lower percentage of sequence identity.
CC This sequence represents plasmid DNA, used in heteroduplex substrate
CC preparation
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 8; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

QY 1 CTTAATTTGTAAGGCTTAATATTTGTTAAATTCGCGTTAATTTGTTAAATCACTC 60
DB 3637 CTTAATTTGTAAGGCTTAATATTTGTTAAATTTGCGTTAATTTGTTAAATCACTC 3578
QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTAATTAATCAAAAGATGACCGA 120
DB 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTAATTAATCAAAAGATGACCGA 3518

QY 121 GATAGGTTGAGTGTGTTCCAGTTTGGACAAAGATCCACTATTAAAGACGTGACATC 180
| | | | |
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGACAAAGATCCACTATTAAAGACGTGACATC 3458
| | | | |
QY 181 CAACGTCAAAAGGCGCAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACATCAAC 240
| | | | |
Db 3457 CAACGTCAAAAGGCGCAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACATCAAC 3398
| | | | |
QY 241 CTAAATCAAGTTTTTTGGGGTGCAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAG 300
| | | | |
Db 3397 CTAAATCAAGTTTTTTGGGGTGCAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAG 3338
| | | | |
QY 301 CCCCCGATTTTAAAGCTTGAACGGGGAAAGCCGGCGAAACGTGGCGAAGAAAGGAAGAA 360
| | | | |
Db 3337 CCCCCGATTTTAAAGCTTGAACGGGGAAAGCCGGCGAAACGTGGCGAAGAAAGGAAGAA 3278
| | | | |
QY 361 AGCGAAAGAGCGGGCGCTAAGGCGCTGGCAAGTGAAGCGGTCAAGCTGCGGCTAACAC 420
| | | | |
Db 3277 AGCGAAAGAGCGGGCGCTAAGGCGCTGGCAAGTGAAGCGGTCAAGCTGCGGCTAACAC 3218
| | | | |
QY 421 CACACCCGCGCGCTTAATGCGCGCTACAGGCGCGTCCATTGCGCAATTGAGGCTGCG 480
| | | | |
Db 3217 CACACCCGCGCGCTTAATGCGCGCTACAGGCGCGTCCATTGCGCAATTGAGGCTGCG 3158
| | | | |
QY 481 CAACGTGTGGGAAGGGCGCATCGGTGCGGGCTCTTTCGTATTACCGCAGCTGGCGAAAG 540
| | | | |
Db 3157 CAACGTGTGGGAAGGGCGCATCGGTGCGGGCTCTTTCGTATTACCGCAGCTGGCGAAAG 3098
| | | | |
QY 541 GGGATGTGCTGCAGGCGCATTAAGTTGGGTAAAGCGAGGGTTTTCCAGTCAAGAGCTTG 600
| | | | |
Db 3097 GGGATGTGCTGCAGGCGCATTAAGTTGGGTAAAGCGAGGGTTTTCCAGTCAAGAGCTTG 3038
| | | | |
QY 601 TAAAAAGACGCGCACTGAGCGCGCTCGTTCAATTCAGCTTTTGAACCCGTGGAGAGAG 660
| | | | |
Db 3037 TAAAAAGACGCGCACTGAGCGCGCTCGTTCAATTCAGCTTTTGAACCCGTGGAGAGAG 2978
| | | | |
QY 661 GCAGACTGCGCGTGCAGAAATGTGTTTTTACGCGTGAATGAGAGATGAAGATCTGCAGAC 720
| | | | |
Db 2977 CCGCGGTGCGCGCTGCAGAAATGTGTTTTTACGCGTGAATGAGATGAAGATCTGCAGAC 2918
| | | | |
QY 721 GCTGAGAAACAGCAGCTAGATTAAACCTAGAAAGATATCATATTGAGCGTAGCTTAA 780
| | | | |
Db 2917 ATAGTTCAATCCATGCCATGTGTAATCCAGCAGCTGTTACAACTCAAGAGAGCAATGT 2858
| | | | |
QY 781 AGATATATC-ATGCGTAAATTTGACGATGAGATCTGTAATACGACTAGTAAAGGCGCA 838
| | | | |
Db 2857 GGTCTCTCTTTTCGTTGGGATCTTTGAAAGGGCAGATTTGTGACAGGTAAATGTTGT 2798
| | | | |
QY 839 ATTGGGTACCGGGCCCCCCTCGAGAGTGAACGTAATCGATTAAGCTTGAATTCGAATTCCT 898
| | | | |
Db 2797 CTGGTAAAGAGCAGAGGCGCATGCGCAATTTGTTGATTAATAGTGTCTGTAATTT 2738
| | | | |
QY 899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCGCGCACCGCGGTGAGCTCCAGCTTT 958
| | | | |
Db 2737 GAGCGCTTCATCTTCAATGTGTGTCTAATTTTAAAGTTAACTTTGATTCATCTTTT 2678
| | | | |
QY 959 GTTCCCTTAAAGAGGTTAATTAATCCATGCGTCAATTTTACGACAGACTATCTTCT 1018
| | | | |
Db 2677 GTTGTGTGCAATGATATATCATTTGTGGA-GTATAGTTGTATTCATTTTGTGTCA 2619
| | | | |
QY 1019 AGGGTAAATCTAGTGCATCAGATCATATGCTGGGTCTTTTTCGGGCTCAGTCAATCG 1078
| | | | |
Db 2618 AGAATGTTTCATCTTCTTAAATCAATACCTTTTAACTCATTC-----TATTAACA 2564
| | | | |
QY 1079 CCCAAGCTGGCGCTATCTGGGCGATGGGGAAGAAAGCCGCTGCTTTTCCCGGAGAGT 1138
| | | | |
Db 2563 GGGTATCACCTTCAACTTGACTTC-----AGCAGTGTCTGTAGTTCCCTC 2515
| | | | |
QY 1139 TGAACGGCGATGAAAGATTTGCGAGAGATGACTGCTGCTGATGAAGTTGAGCGAA 1198
| | | | |
Db 2514 ATCTTGAATAATATATGTTCTTCTGTATACATTAACCTTCGGGCAATGGCACTCTTGA 2455
| | | | |

QY 1199 ACGCAGTTTACATGATGATTCGGGAAGTGTGCGCATGACCGCTTAAACGTGAATC 1258
| | | | |
Db 2454 GTCATGCCGTTTCAATATGATCTGGGTATCTTGAAGAAAGATTGAACCATTAAGAAAGT 2395
| | | | |
QY 1259 GTTCCGTAGGCGCACCTGGGATTAACAGTTCGTGCGCGCTTTTCCGGACACAGTTCCGAT 1318
| | | | |
Db 2394 AGTGAAGTGTGGCCATGGAACAGATAGTTTTCAGTAACTAGTCAAAATTAATTAAGGT 2335
| | | | |
QY 1319 GGTCAAGCCCGAAGCGCATCAGCAACCCGAAACATACCGGCAAGCGGAACTGCGGTGC 1378
| | | | |
Db 2334 AAGTTTTC-----CGTATGTTGATCACTTCAACCTCTCCACTGACAGAAATTTGTGC 2280
| | | | |
QY 1379 CGGTGTGAGATTAATGACAGCGGTGCGCGCTGGGAATTAACGTACAGCGAGAGGGTA 1438
| | | | |
Db 2279 CCATTAACATCAACATCTAATTTCAAGAAATGGGACAACTCCAGTAAAGATCTTCT 2220
| | | | |
QY 1439 TCTGTGCTGATGCGCGAAGATGAACATGATACCCGTGATTAACCGGCGGGCGGC 1498
| | | | |
Db 2219 CTTTACTCATTCGGTACCCAGCTTTTGTCCCTTTAGTGAAGGTTA-----ATTGCGGC 2165
| | | | |
QY 1499 TTGGGTAATCATGATCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCAAAATTCA 1558
| | | | |
Db 2164 TTGGGTAATCATGATCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCAAAATTCA 2105
| | | | |
QY 1559 CACAAATACAGAGCCGGAAGCATTAAGTGTAAAGCTTGGGGTGCCTAATGATGACTTA 1618
| | | | |
Db 2104 CACAAATACAGAGCCGGAAGCATTAAGTGTAAAGCTTGGGGTGCCTAATGATGACTTA 2045
| | | | |
QY 1619 CTCAATTAATTTGGGTTGCGCTCACTGCGCGCTTTCCAGTGGGGAACCTGTGCGGACG 1678
| | | | |
Db 2044 CTCAATTAATTTGGGTTGCGCTCACTGCGCGCTTTCCAGTGGGGAACCTGTGCGGACG 1985
| | | | |
QY 1679 CTGATTAATGAATCGGCAACCGCGGGAGAGCGGTTTGCATATTGGCGCTCTTCC 1738
| | | | |
Db 1984 CTGATTAATGAATCGGCAACCGCGGGAGAGCGGTTTGCATATTGGCGCTCTTCC 1925
| | | | |
QY 1739 GCTTCTCTGCTCACTGACTCGCTGCGTGGTCTTGGGCTGGCGGAGCGGTATCACT 1798
| | | | |
Db 1924 GCTTCTCTGCTCACTGACTCGCTGCGTGGTCTTGGGCTGGCGGAGCGGTATCACT 1865
| | | | |
QY 1799 CACTCAAAAGCGGTATATCGGTTATCCAGAAATCAGGGGATTAACGAGAAAGAAACATG 1858
| | | | |
Db 1864 CACTCAAAAGCGGTATATCGGTTATCCAGAAATCAGGGGATTAACGAGAAAGAAACATG 1805
| | | | |
QY 1859 TGAGCAAAAGGCGCAGCAAAAGCGCAAGACCGTAAAGAGCGCGTGTGCGCTTTTC 1918
| | | | |
Db 1804 TGAGCAAAAGGCGCAGCAAAAGCGCAAGACCGTAAAGAGCGCGTGTGCGCTTTTC 1745
| | | | |
QY 1919 CATAGGCTCGCGCCCTGACGACATCAAAAATGAGCGCTCAAGTCAAGGTGCGCA 1978
| | | | |
Db 1744 CATAGGCTCGCGCCCTGACGACATCAAAAATGAGCGCTCAAGTCAAGGTGCGCA 1685
| | | | |
QY 1979 AACCGACAGGACTATAAGATACAGGCGTTTCCCTGGAAGTCCCTGATCGCTCT 2038
| | | | |
Db 1864 AACCGACAGGACTATAAGATACAGGCGTTTCCCTGGAAGTCCCTGATCGCTCT 1625
| | | | |
QY 2039 CCTGTTCGACCTTCGCGCTTACCGGATTAACGTGCTGCGCTTCTCCCTTCCGGAAGCGTG 2098
| | | | |
Db 1624 CCTGTTCGACCTTCGCGCTTACCGGATTAACGTGCTGCGCTTCTCCCTTCCGGAAGCGTG 1565
| | | | |
QY 2099 GCGCTTCTCATAGCTCAAGCTGTAGATATCTCAGTTGCGGTAGTGTGCTTCCAG 2158
| | | | |
Db 1564 GCGCTTCTCATAGCTCAAGCTGTAGATATCTCAGTTGCGGTAGTGTGCTTCCAG 1505
| | | | |
QY 2159 CTGGGCTGTGACAGAAACCCCGCTCAGCGGACCGGCTGCGCTTATCCGTAATCTAT 2218
| | | | |
Db 1504 CTGGGCTGTGACAGAAACCCCGCTCAGCGGACCGGCTGCGCTTATCCGTAATCTAT 1445
| | | | |
QY 2219 CGTCTTGAATCAACCCGATTAAGACAGACTTATCGCACTGCGAGAGCACTGTATAC 2278
| | | | |
Db 1444 CGTCTTGAATCAACCCGATTAAGACAGACTTATCGCACTGCGAGAGCACTGTATAC 1385
| | | | |
QY 2279 AGGATTAAGCAGAGGAGATATGAGCGGTCTACAGAGTTCTTGAAGTGTGCGCTAAC 2338
| | | | |

```

Db      1384 AGGATTAGACGACGAGGATGATGAGCGGATGCTACAGAGTTCTTGAAGTGATGGCTTAAC 1325
Qy      2339 TAGGCTACCTAGAAAGACAGTATTTGGTATCTGGGCTCTGCGAAGCCGTTAACTTC 2398
Db      1324 TAGGGCTACCTAGAAAGACAGTATTTGGTATCTGGGCTCTGCGAAGCCGTTAACTTC 1265
Qy      2399 GGAAGAAAGTTGGTATGCTCTGATCCGCAACCAACCGCTGTAGCGGTGTTT 2458
Db      1264 GGAAGAAAGTTGGTATGCTCTGATCCGCAACCAACCGCTGTAGCGGTGTTT 1205
Qy      2459 TTTGTTTGCAAGCAGACAGATTACGCGGAAAAAAGGATCTCAAGAAAGTCTTTGATC 2518
Db      1204 TTTGTTTGCAAGCAGACAGATTACGCGGAAAAAAGGATCTCAAGAAAGTCTTTGATC 1145
Qy      2519 TTTTCTACGGGGTCTGAGCTCAGTGGAAAGAAACTCACGTTAAGGGATTTTGTGATG 2578
Db      1144 TTTTCTACGGGGTCTGAGCTCAGTGGAAAGAAACTCACGTTAAGGGATTTTGTGATG 1085
Qy      2579 AGATTATCAAAAAGGATCTTCACTAGATCCTTTTAATTAATAAGTTTAAATCA 2638
Db      1084 AGATTATCAAAAAGGATCTTCACTAGATCCTTTTAATTAATAAGTTTAAATCA 1025
Qy      2639 ATCTAAAGTATATATGATTAACCTTGGCTGACAGTTACCAATGCTTAATCAGTAGGCA 2698
Db      1024 ATCTAAAGTATATATGATTAACCTTGGCTGACAGTTACCAATGCTTAATCAGTAGGCA 965
Qy      2699 CCTATCTACGCGATCTGTCTATTTCTGTTCACTCATGTTGCGTACTCCCGTCTGTAG 2758
Db      964 CCTATCTACGCGATCTGTCTATTTCTGTTCACTCATGTTGCGTACTCCCGTCTGTAG 905
Qy      2759 ATAACTACGATAGGGAGGGCTTAACATTTGGCCCAAGTGTGCAATGATACCGGAGAC 2818
Db      904 ATAACTACGATAGGGAGGGCTTAACATTTGGCCCAAGTGTGCAATGATACCGGAGAC 845
Qy      2819 CCAGCTCACCGGCTCCAGATTATACAGATTAACCGACCGCCGGAAGGCGGAGCGC 2878
Db      844 CCAGCTCACCGGCTCCAGATTATACAGATTAACCGACCGCCGGAAGGCGGAGCGC 785
Qy      2879 AGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCT 2938
Db      784 AGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCT 725
Qy      2939 AGAGTATGATGTTCCAGTTAATAGTTTGGCAACGTTGGCATTTGCTAAGGATC 2998
Db      724 AGAGTATGATGTTCCAGTTAATAGTTTGGCAACGTTGGCATTTGCTAAGGATC 665
Qy      2999 GTGATGTACCGCTCGTCTTGGTATGCTTCAATTCAGTCCGGTCCCAAGATCAAG 3058
Db      664 GTGATGTACCGCTCGTCTTGGTATGCTTCAATTCAGTCCGGTCCCAAGATCAAG 605
Qy      3059 CGAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGCTCTTCCGTCCTCGATC 3118
Db      604 CGAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGCTCTTCCGTCCTCGATC 545
Qy      3119 GTTGTGCAAGTATGTTGGCGGAGTGTATCACTAGTATGSCAGACATGATAT 3176
Db      544 GTTGTGCAAGTATGTTGGCGGAGTGTATCACTAGTATGSCAGACATGATAT 485
Qy      3179 TCTTTAATGTCATGCGCATCCGTAAGATGTTTTCTGACGTAAGTACTCAACCAAG 3238
Db      484 TCTTTAATGTCATGCGCATCCGTAAGATGTTTTCTGACGTAAGTACTCAACCAAG 425
Qy      3239 TCATTTGAGAATATGTTATGCGGCAACGAGTTGCTTTGCGCGCGCAATTCGGGAT 3298
Db      424 TCATTTGAGAATATGTTATGCGGCAACGAGTTGCTTTGCGCGCGCAATTCGGGAT 365
Qy      3299 AATATCGCGCCACATAGCAGAACTTTAAAGGTCATCATTTGAAAAGCTTTCCGGG 3358
Db      364 AATATCGCGCCACATAGCAGAACTTTAAAGGTCATCATTTGAAAAGCTTTCCGGG 305
Qy      3359 CGAAACTCTCAGAGATCTTACCGCTGTGAGATCAGTTGATGTAACCACTCCGTCGA 3418

```

```

Db      304 CGAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTGATTAACCACTCGTNGCA 245
Qy      3419 CCCAATGATCTTACGATCTTTTATCTTTACACAGCGTTTCTGGTGAGCAAAAACAGA 3478
Db      244 CCCAATGATCTTACGATCTTTTATCTTTACACAGCGTTTCTGGTGAGCAAAAACAGA 185
Qy      3479 AGGCAAAATGCGCAAAAAGGGAATPAGGGGCGACACGGAATGTTGAATACTCATCTC 3538
Db      184 AGGCAAAATGCGCAAAAAGGGAATPAGGGGCGACACGGAATGTTGAATACTCATCTC 125
Qy      3539 TTTCTTTTGAATTTTATGACATTTATACAGGTTATTTGTCATGAGCGGATACATA 3598
Db      124 TTTCTTTTGAATTTTATGACATTTATGACGTTATTTGTCATGAGCGGATACATA 65
Qy      3599 TTTGAATGATTTTAAAGAAATTAACAAATPAGGGGTTCCGCGCAATTTCCGAAAGTG 3658
Db      64 TTTGAATGATTTTAAAGAAATTAACAAATPAGGGGTTCCGCGCAATTTCCGAAAGTG 5
Qy      3659 CCAC 3662
Db      4 CCAC 1

```

```

RESULT 8
ADM68462/c
ID ADM68462 standard; DNA; 3637 BP.
AC
XX
AC ADM68462;
XX
DT 03-JUN-2004 (first entry)
XX
XX
XX
XX Wild type green fluorescent protein, GFP, gene construct.
XX
XX
XX ds; mismatch endonuclease; endonuclease; gene shuffling technology;
XX single nucleotide polymorphism; cancer susceptibility;
XX sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
XX Aequorea victoria.
XX Synthetic.
XX
XX US2003157682-A1.
XX
XX
XX PD 21-AUG-2003.
XX
XX
XX PF 31-JAN-2003; 2003US-00356708.
XX
XX
XX PR 01-FEB-2002; 2002US-0353722P.
XX
XX PR 14-MAR-2002; 2002US-00098155.
XX
XX PR 01-AUG-2002; 2002US-00211079.
XX
XX
XX (PADG/) PADGETT H S.
XX (VAEW/) VAENHONGS A A.
XX (VOUT/) VOUTANI F S.
XX (SMIT/) SMITH M L.
XX (LIND/) LINDBO J A.
XX (FITZ/) FITZMAURICE W P.
XX
XX Padgett HS, Vaenhongs AA, Vojdani FS, Smith ML, Lindbo JA;
XX Fitzmaurice WP;
XX
XX WPI; 2003-766176/72.
XX
XX
XX Making a mismatch endonuclease, useful in gene shuffling and in detection
XX of single nucleotide polymorphisms, comprises transfecting a host with a
XX recombinant viral vector including a polynucleotide encoding a mismatch
XX endonuclease.
XX
XX Example 5; SEQ ID NO 31; 79pp; English.
XX
XX The invention relates to a method of making a mismatch endonuclease
XX enzyme comprising transfecting a host plant, animal, yeast, fungus or
XX bacterium with a recombinant viral vector that encodes a polynucleotide
XX sequence for a mismatch endonuclease, growing the host so that the

```


CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for
CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents the wild type green fluorescent protein, GFP,
CC gene construct.

SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 11; Length 3637;

Best Local Similarity 84.0%; Pred. No. 0;

Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

```
OY 1 CTAATTGTAAAGCGTTAATATTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3637 CTAATTGTAAAGCGTTAATATTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
OY 61 ATTTTAAACCAATAGCCGGAATCGCAAAATCCCTTAATAATCAAAAGATAGACGA 120
DB 3577 ATTTTAAACCAATAGCCGGAATCGCAAAATCCCTTAATAATCAAAAGATAGACGA 3518
OY 121 GATAGGTTGAGTGTGTTCCAGTTTGAAACAAGTCCACTATTAAAGAGTGCATC 180
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGAAACAAGTCCACTATTAAAGAGTGCATC 3458
OY 181 CAACGTCAAAGGCGCAAAACCGTCTATACAGGCGAGTGGCCACTACGTAACATCAAC 240
DB 3457 CAACGTCAAAGGCGCAAAACCGTCTATACAGGCGAGTGGCCACTACGTAACATCAAC 3398
OY 241 CTAATCAAGTTTTTGGGCGTGAAGTCCGTAAGCACTAAATCGAAACCTTAAGGAG 300
DB 3397 CTAATCAAGTTTTTGGGCGTGAAGTCCGTAAGCACTAAATCGAAACCTTAAGGAG 3338
OY 301 CCCCCGATTTAGAGCTTTGACGGGAAAGCCGCGCAACGTGGCGAAGAAAGGAGAA 360
DB 3337 CCCCCGATTTAGAGCTTTGACGGGAAAGCCGCGCAACGTGGCGAAGAAAGGAGAA 3278
OY 361 AGCGAAAGAGCGGCGCTGAGGCGCTGACAAAGTAGCGGTCAAGCTGCGCGTAACAC 420
DB 3277 AGCGAAAGAGCGGCGCTGAGGCGCTGACAAAGTAGCGGTCAAGCTGCGCGTAACAC 3218
OY 421 CACACCGCGCGCTTAAATGCGCGCTACAGGCGCGTCCCATTTGCCATTCAAGCTCG 480
DB 3217 CACACCGCGCGCTTAAATGCGCGCTACAGGCGCGTCCCATTTGCCATTCAAGCTCG 3158
OY 481 CAACGTGTGGAAAGGCGCATCGTTCGCGCTTTCCGTAATTAAGCCAGCTGGCGAAG 540
DB 3157 CAACGTGTGGAAAGGCGCATCGTTCGCGCTTTCCGTAATTAAGCCAGCTGGCGAAG 3098
OY 541 GGGATGTCTGCAAGCGCATTAAGTTGGGTAAACGCGAGGGTTTCCAGTACAGCGTTG 600
DB 3097 GGGATGTCTGCAAGCGCATTAAGTTGGGTAAACGCGAGGGTTTCCAGTACAGCGTTG 3038
OY 601 TAAACGACGCGCACTGACGCGCTCGTTCATTCACGTTTTTGAACCCGTGAGAGACG 660
DB 3037 TAAACGACGCGCACTGACGCGCTCGTTCATTCACGTTTTTGAACCCGTGAGAGACG 2978
OY 661 GCAAGCTCGCGGTGCAAAATGTGTTTAAAGCGTGAATGAGACATGCTGCACAC 720
DB 2977 CCGCGGTGCGCGGTGCAAAATGTGTTTAAAGCGTGAATGAGACATGCTGCACAC 2918
OY 721 GCTCAGAACACGAGCTAGATTAACTTAGAAAGATATATCATATTGAGAGTAACTTAA 780
DB 2917 ATAGTTCATCATGCAATGTGTAATCCAGCAGCTGTTACAACTCAAGAGAGACATGT 2858
OY 781 AGATTAATC-ATGCGTAAATTTGACGATGGGATCTGTAAATACGACTAATAGGCGA 838
DB 2857 GGTCTCTCTTTTCTGTTGGGATCTTTTCAAAAGGCGAGATTGTGTGACAGGTAATGTTGT 2798
```

```
OY 839 ATTGCGTACCGGGCCCCCTCGAGTCGACGTAATCATATGCTGAATTCATTCCT 898
DB 2797 CTGTAAAGAGACAGGGCGCATCGCAATGAGATATTTGTGTAATATGCTGCTAGTT 2738
OY 899 GCAGCCCGGGGATATCACTAGTTCTAGACGGCGCCACCGCGGTGAGCTCCAGCTTTT 958
DB 2737 GAGCGCTTCATCTTCAATGTGTGTCTAATTTTGAAGTAACTTTGATTCATTTCTTTT 2678
OY 959 GTTCCCTTAAAGAGGTTAATTAAGTCCATGCGTCAATTTTACGACACTATCTTTCT 1018
DB 2677 GTTGTCTGCGCATGATGATATCATTTGTGTGA-GTTATGTGTATTCATTTGTGTCA 2619
OY 1019 AGGTTAATCTAGCTGATCAAGATCATATGTGGGCTTTTTCGGCTCAGTATCG 1078
DB 2618 AGAATGTTTCATCTTTTAAATCAATACCTTTTAACTCGATTC-----TATTAAACA 2564
OY 1079 CCGAAGCTGGCGCTATCTGGGCGATCGGGGAGAAAGACCCGTGCTTTTCCCGGAGAGT 1138
DB 2563 GGGTATCACTTCAAACTTGACTTC-----AGCAGTGTCTTGTAGTTCCGTC 2515
OY 1139 TGAAGCGCATGAAAGAGTTTGCAGAGATGACTGCTGCATTTGACGTTGACGAAA 1198
DB 2514 ATCTTGAATAATATATGTTCTTTCTGTATACATTAACCTTCGGGCAATGGCACTTGAAA 2455
OY 1199 ACGACGTTTACATGATGATTTGGGAAAGTGTGGCCATGCAACGCTTTAACGTTGA 1258
DB 2454 GTCATGCCGTTTCAATGATCTGGGTATCTTGAAGAAAGCATTTGAACACCATTAAGAAAGT 2395
OY 1259 GTTCGTTACAGCCACCTGGGATACCAAGTTCGTCGCGCTTTTCCGGAACACAGTTCCG 1318
DB 2394 AGTACAAAGTTTGGCCATGGAACAGTATGTTTCCAGTATGCAATTAATTAAGGAT 2335
OY 1319 GGTACGCCGAAGCGCATACAGCAACCGAACATACCGGCAACAGCCGGAACCTGCGTC 1378
DB 2334 AAGTTTTC-----GTATGTTGCAATCACTTCAACCTCTCCACGACGAAATTTGTGC 2280
OY 1379 CCGTGTGCAATTAATGACAGCGGTGGCGCTGGGATTTAAGTCAGCGAGACGGGTA 1438
DB 2279 CCATTAAACATGACCATCTTAATTTCAACAAAGATTTGGGCACTCCAGTGAAGAAATGTTCTT 2220
OY 1439 TCCTGGCGTGAATGCGCGAAGATGACATGGAATACCCCGTGAATTAACCGCGGGCGCGC 1498
DB 2219 CTTTATCTCATCGGTACCCAGCTTTTGTTCCTTTAATGAGAGGTTA-----ATTGCGCGC 2165
OY 1499 TTGGCGTAAATCATGATCATAGCTGTTTCTGTGTGAATTTGTAATTCGCTCACAATTC 1558
DB 2164 TTGGCGTAAATCATGATCATAGCTGTTTCTGTGTGAATTTGTAATTCGCTCACAATTC 2105
OY 1559 CACAAATATGAGCGCGGAAGATTAAGTGAAGCTTGGGGTGCCTTAATGATGAGCTTA 1618
DB 2104 CACAAATATGAGCGCGGAAGATTAAGTGAAGCTTGGGGTGCCTTAATGATGAGCTTA 2045
OY 1619 CTCAATTAAATTTGGTGTGCGCTCACTGCGCGCTTTCCAGTGGGAAACCTTGCNGCAG 1678
DB 2044 CTCAATTAAATTTGGTGTGCGCTCACTGCGCGCTTTCCAGTGGGAAACCTTGCNGCAG 1985
OY 1679 CTGCATTAAATGATTCGCGCAACGCGCGGAGAGACGAGTTTGCATTTTGGCGCTTTCC 1738
DB 1984 CTGCATTAAATGATTCGCGCAACGCGCGGAGAGACGAGTTTGCATTTTGGCGCTTTCC 1925
OY 1739 GCTTCCTCGCTCATCTGACTGCTGCGCTCGGTGTTTGGGCTGGCGGACGGTATCACT 1798
DB 1924 GCTTCCTCGCTCATCTGACTGCTGCGCTCGGTGTTTGGGCTGGCGGACGGTATCACT 1865
OY 1799 CACTCAAGGCGGTAATACGGTTATTCACAGAAATCAGGAGGTTAAACGAGGAAAGAAATG 1858
DB 1864 CACTCAAGGCGGTAATACGGTTATTCACAGAAATCAGGAGGTTAAACGAGGAAAGAAATG 1805
OY 1859 TGAGCAAAAGCGCAGCAAAAGGCGAGAACCTGTAAGAAAGCGCGTGTGCGCTTTTTC 1918
DB 1804 TGAGCAAAAGCGCAGCAAAAGGCGAGAACCTGTAAGAAAGCGCGTGTGCGCTTTTTC 1745
OY 1919 CATAGGCTCCGCCCCCTGACGAGCATCAAAAAATGACGCTCAAGTCAAGGTGCGCA 1978
```

Db	1744	CATGAGCTCCGCCCCCTGACGAGCATCAAAATATGACGCTCAAGTCAGAGGTGGGA	1685
QY	1979	AAACCGACAAGACATTAAGAATACACAGCGCTTTCCCTGGAAGCTCCTGTCGCTCT	2038
Db	1684	AACCGACAAGACATTAAGAATACACAGCGCTTTCCCTGGAAGCTCCTGTCGCTCT	1625
QY	2039	CCTGTTCCGACCCCTCGGCTTACCGATACCTGTCCGCTTCTCCCTCGGGAAGCGTG	2098
Db	1624	CTGTGTCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTG	1565
QY	2099	GCGCTTCTCATAGCTACCGCTGAGGATCTCAGTTCCGTGTAGCTGTTGCTCAAG	2158
Db	1564	GCGCTTCTCATAGCTACCGCTGAGGATCTCAGTTGCTGTAGCTGTTGCTCAAG	1505
QY	2159	CTGGGCTGTGTGCAAGAACCCCGTTACGCCGACCGGTGGCGCTTATCCGGTAACTAT	2218
Db	1504	CTGGGCTGTGTGCAAGAACCCCGTTACGCCGACCGGTGGCGCTTATCCGGTAACTAT	1445
QY	2219	CGTCTTAGTCCAAACCCGGTAAACACGACTTATGCGCATCTGGACAGCCACTGTATAC	2278
Db	1444	CGTCTTAGTCCAAACCCGGTAAACACGACTTATGCGCATCTGGACAGCCACTGTATAC	1385
QY	2279	AGGATTAGCAGACGAGGTATGTAGCGCGGTGTACAGAGTCTTGAAAGTGATGGCTTAC	2338
Db	1384	AGGATTAGCAGACGAGGTATGTAGCGCGGTGTACAGAGTCTTGAAAGTGATGGCTTAC	1325
QY	2339	TACGGCTACACTGAAGAGCAGATTTTGGTATCTGCGTCTGTGAAGCCAGTTTACCTTC	2398
Db	1324	TACGGCTACACTGAAGAGCAGATTTTGGTATCTGCGTCTGTGAAGCCAGTTTACCTTC	1265
QY	2399	GGAAAAAGATTGTAGCTCTTGATCCGGCAAAACAAACCGGTGGTAGCGGTGTTT	2458
Db	1264	GGAAAAAGATTGTAGCTCTTGATCCGGCAAAACAAACCGGTGGTAGCGGTGTTT	1205
QY	2459	TTTGTTTGCACAGACAGATTACCGGCGCAAAAAAGATCTCAAGAGATCTTTGATC	2518
Db	1204	TTTGTTTGCACAGACAGATTACCGGCGCAAAAAAGATCTCAAGAGATCTTTGATC	1145
QY	2519	TTTTTCTACGGGCTCTGACGCTCAGTGTGAGGAAACCTCAGTAAAGGATTTTGGTCATG	2578
Db	1144	TTTTTCTACGGGCTCTGACGCTCAGTGTGAGGAAACCTCAGTAAAGGATTTTGGTCATG	1085
QY	2579	AGATTATCAAAAAAGATCTTCACTTAATCTTTTAAATTTAAATGAGTTTAAATCA	2638
Db	1084	AGATTATCAAAAAAGATCTTCACTTAATCTTTTAAATTTAAATGAGTTTAAATCA	1025
QY	2639	ATCTAAATATATATGATTAACCTTGCTGTGACAGTTTACATGCTTATACATGAGGCA	2698
Db	1024	ATCTAAATATATATGATTAACCTTGCTGTGACAGTTTACATGCTTATACATGAGGCA	965
QY	2699	CCTATCTCAGGATCTGTCTAATTTGTTCAATCATATGTTGCTGTGCTCCCGTGTAG	2758
Db	964	CCTATCTCAGGATCTGTCTAATTTGTTCAATCATATGTTGCTGTGCTCCCGTGTAG	905
QY	2759	ATTAACCTACGATCGGAGGCGCTTACCATCTGCGCCACAGTGTGCAATGATTCGCGGAC	2818
Db	904	ATTAACCTACGATCGGAGGCGCTTACCATCTGCGCCACAGTGTGCAATGATTCGCGGAC	845
QY	2819	CCAGGCTCACCGGCTCAATTTATACGATTAACCGACCGACCGGAAGGCGCGAGCG	2878
Db	844	CCAGGCTCACCGGCTCAATTTATACGATTAACCGACCGACCGGAAGGCGCGAGCG	785
QY	2879	AGAAAGTGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCGGGAAGCT	2938
Db	784	AGAAAGTGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCGGGAAGCT	725
QY	2939	AGAGTAAGATGTTCCGCAATTAATGTTTGTGCGAAGCTGTTGCAATGCTACAGGATC	2998
Db	724	AGAGTAAGATGTTCCGCAATTAATGTTTGTGCGAAGCTGTTGCAATGCTACAGGATC	665
QY	2999	GTTGTTGTCACGCTCGTCTGTTGGTATGCTTCAATTCAGCTCCGTTCCACGATCAAG	3058

Db	664	GTGGTGTCACGCTCGTGGTGGTATGGCTTCATTCAAGCTCCGGTTCCCAACATCAAG	605
Qy	3059	CGAGTTACATGATGCCCATGTGGTGCAAAAACGGTTAGCTCTTCGGTCTCCGATC	3111
Db	604	CGAGTTACATGATGCCCATGTGGTGCAAAAACGGTTAGCTCTTCGGTCTCCGATC	545
Qy	3119	GTTCGACAAAGTAAATTGGCCGCGAGTGTATCACTCAATGGTTATGGCAGCATGCATTAAT	317
Db	544	GTTCGACAAAGTAAATTGGCCGCGAGTGTATCACTCAATGGTTATGGCAGCATGCATTAAT	485
Qy	3179	TCCTTACTGTTCATGCCATCCGTAAGATGCTTTCTGTGTCTGTGAATAGTAACTCAACAAAG	323
Db	484	TCCTTACTGTTCATGCCATCCGTAAGATGCTTTCTGTGTCTGTGAATAGTAACTCAACAAAG	425
Qy	3239	TCATTCTGGAATAGTGTATTCGGCGCACCGAGTTGCTTTGCCCCGCGCTCAATACCGAGAT	3298
Db	424	TCATTCTGGAATAGTGTATTCGGCGCACCGAGTTGCTTTGCCCCGCGCTCAATACCGAGAT	365
Qy	3299	AATACCGCGCCATACAGAACTTTAAAGTGTGTCATCATTTGGAAAAAGCTTTCGGGG	335
Db	364	AATACCGCGCCATACAGAACTTTAAAGTGTGTCATCATTTGGAAAAAGCTTTCGGGG	305
Qy	3359	CGAAATCTCAAGAGATCTTACCGCTGTGAATCCAGTTCGATGTAAACCACTCGTGCA	3418
Db	304	CGAAATCTCAAGAGATCTTACCGCTGTGAATCCAGTTCGATGTAAACCACTCGTGCA	245
Qy	3419	CCCAACTGATCTTCAGCATCTTTTACTTTCACACAGCGTTCTGGGTGAGCAAAAACAGGA	3478
Db	244	CCCAACTGATCTTCAGCATCTTTTACTTTCACACAGCGTTCTGGGTGAGCAAAAACAGGA	185
Qy	3479	AGGCAAAAAGCCGCAAAAAGGAATTAAGGGGAGACGAAATGTTGAATCTCATATC	3538
Db	184	AGGCAAAAAGCCGCAAAAAGGAATTAAGGGGAGACGAAATGTTGAATCTCATATC	125
Qy	3539	TTCTCTTTTCAATATTATTGGAAGCATTTATCAGGGTTATGTCTCATGAGCGAATACATA	3598
Db	124	TTCTCTTTTCAATATTATTGGAAGCATTTATCAGGGTTATGTCTCATGAGCGAATACATA	65
Qy	3599	TTTGATGTATTGAAAAATTAACAATATAGGGGTTCCGGCACATTTCCCGAAAAAGTG	3658
Db	64	TTTGATGTATTGAAAAATTAACAATATAGGGGTTCCGGCACATTTCCCGAAAAAGTG	5
Qy	3659	CCAC 3662	
Db	4	CCAC 1	
RESULT 9			
ADP26615/c			
ID	ADP26615	standard; DNA; 3637 BP.	
XX	AC	ADP26615;	
XX	XX		
DT	26-AUG-2004	(first entry)	
XX	DE	Green fluorescent protein (GFP) plasmid DNA #1.	
XX	KW	Sequence variation; heteroduplex; transcription; DNA integration;	
XX	XX	ribozyme expression; gene; de; green fluorescent protein; GFP.	
OS	Aequorea victoria.		
XX	OS	Synthetic.	
XX	PN	US2004110130-A1.	
XX	PD	10-JUN-2004.	
XX	PF	25-OCT-2002; 2002US-00280913.	
XX	PR	02-FEB-2001; 2001US-0266386P.	
XX	PR	14-FEB-2001; 2001US-0268785P.	
XX	PR	01-FEB-2002; 2002US-0006390.	
XX	PR	08-AUG-2002; 2002US-0402342P.	

XX (LARG-) LARGE SCALE BIOLOGY CORP.
 XX Padgett HS, Lindbo JA, Fitzmaurice WP;
 XX WPI; 2004-440326/41.
 XX
 PT Redistributing sequence variations between non-identical polynucleotide
 PT sequences, useful for generating improved polynucleotide having a desired
 PT characteristic, comprises making a heteroduplex and introducing a nick.
 XX
 PS Example 5; SEQ ID NO 31; 75bp; English.
 XX
 CC The invention relates to an in vitro method of redistributing sequence
 CC variations between non-identical polynucleotide sequences, comprising
 CC making a heteroduplex polynucleotide from two non-identical
 CC polynucleotides, introducing a nick in the second strand at or near a
 CC base pair mismatch site, removing the mismatched base(s) from the
 CC mismatch site where the nick occurred and using the first strand as a
 CC template to replace the removed base(s) with bases that complement the
 CC base(s) in the first strand. The invention also relates to an in vitro
 CC method of making a population of sequence variants from a heteroduplex
 CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
 CC desired functional property and identifying a reassorted DNA molecule
 CC encoding a protein with a desired functional property. The method is
 CC useful for generating an improved polynucleotide sequence or a population
 CC of improved polynucleotide sequences possessing at least one desired
 CC phenotypic characteristic (e.g., promotes transcription of linked
 CC polynucleotides), where such polynucleotides are useful for expression
 CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
 CC integration to form a transgenic plant, animal or microorganism, and for
 CC expression of a ribozyme. This sequence represents DNA used in the scope
 CC of the invention.
 XX
 SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 12; Length 3637;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

QY 1 CTAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
 DB 3637 CTAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
 QY 61 ATTTTAAACCAATAGCGCGAATCGCAAAATCCCTTAATATCAAAAGATACCGGA 120
 DB 3577 ATTTTAAACCAATAGCGCGAATCGCGAATCCCTTAATATCAAAAGATACCGGA 3518
 QY 121 GATAGGGTGAAGTGTTCAGTTTGAACAAAGTCCACTAATTAAGAAAGCGGACTC 180
 DB 3517 GATAGGGTGAAGTGTTCAGTTTGAACAAAGTCCACTAATTAAGAAAGCGGACTC 3458
 QY 181 CAACGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGCGCCACTAGTGAACCATCAC 240
 DB 3457 CAACGTCAAAAGGCGGAAACCGTCTATCAGGGCGATGCGCCACTAGTGAACCATCAC 3398
 QY 241 CTAAATCAAGTTTGTGGGCTGAGGTGCGCTTAAGCACTAAATCGGAACCTTAAGGAG 300
 DB 3397 CTAAATCAAGTTTGTGGGCTGAGGTGCGCTTAAGCACTAAATCGGAACCTTAAGGAG 3338
 QY 301 CCCCCGATTAGAGCTTGACGGGGGAAAGCCGGCGCAAGTGGCGGAAAGGAAGGAAGA 360
 DB 3337 CCCCCGATTAGAGCTTGACGGGGGAAAGCCGGCGCAAGTGGCGGAAAGGAAGGAAGA 3278
 QY 361 AGGGAAGGAGCGGGCGCTAGGGCGCTGCAAGTGAAGCGGTCAAGCTGCGCTGAACAC 420
 DB 3277 AGGGAAGGAGCGGGCGCTAGGGCGCTGCAAGTGAAGCGGTCAAGCTGCGCTGAACAC 3218
 QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGCTCCATTGCGCATTCAGGCTGCG 480
 DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGCTCCATTGCGCATTCAGGCTGCG 3158
 QY 481 CAACTGTTGGGAAGGGGATCGGTGCGGGCCTCTTCGCTATTAGCCAGCTGGGGAAGG 540

DB 3157 CAACTGTTGGGAAGGGGATCGGTGCGGGCCTCTTCGCTATTAGCCAGCTGGGGAAGG 3098
 QY 541 GGAATGTCCTGCAAGCGGATTAAGTGGTAAACGCCAGGGTTTCCAGTCAAGACGTTG 600
 DB 3097 GGAATGTCCTGCAAGCGGATTAAGTGGTAAACGCCAGGGTTTCCAGTCAAGACGTTG 3038
 QY 601 TAAACGACGGCCAGTGAAGCGCGCTGCTTCAATTCAGGTTTAAACCCCGTGAAGACG 660
 DB 3037 TAAACGACGGCCAGTGAAGCGCGCTGCTTCAATTCAGGTTTAAACCCCGTGAAGACG 2978
 QY 661 GCACAGCTCGGGTCAATGTTTACAGGTGATGAGCAAGTGAAGATGCTGCACAC 720
 DB 2977 CCGGCGTGGGGCGCGCTCTGAAGTAAGTGAATCCCCGGCTGCAGGAATTTCTTATTTGT 2918
 QY 721 GCTGCAGAACACGACGATTAATTAACCTTAAGAAAGATATCATATTTGACGTAA 780
 DB 2917 ATAGTATCATCATGATGATGATATCCACAGAGCTTTACAACTCAAGAAAGACATGT 2858
 QY 781 AGATATTC--ATGCGTAAATTTGACGATGGATCTGTATTAAGACTCATTATGGCGCA 838
 DB 2857 GGTCTCTTTTGTGGGATCTTTCCAAAGGGCAGATGTGTGACAGGTAATGTGTGT 2798
 QY 839 ATTGGATACGGGGCCCCCGTCAAGTGAAGCGGATGATTAAGCTTAATTCGAAATTCCT 898
 DB 2797 CTGCTAAAGGACAGGGCCATCCCAATTGGAATTTTGTGTAATGGTCTGCTAGT 2738
 QY 899 GCAAGCCCGGGGATTCACATGTTCTAGAGGGCGCCACCGCGGTGAGCTCCAGCTTTT 958
 DB 2737 GAACGCTTCATCTCAATGTTGTGTCTAATTTTGAAGTTAATCTTATTCATTTCTTTT 2678
 QY 959 GTTCCCTTATGATGAGGTTAATTAATCCATGCGTCAATTTTACGACATCTTTCT 1018
 DB 2677 GTTGTCTGCCAGTATGATTAATTAATGTTGTA--GTATAGTGTATTAATTCATTTGTGCA 2619
 QY 1019 AGGTTATATCAGTGCATAGAGATCATATCGTCGGCTTTTTCGGGCTGATCATG 1078
 DB 2618 AGAATGTTTCATCTTTTAAATTAATATCTTTTAATCTGATTC-----TATTAACA 2564
 QY 1079 CCCAAGCTGCGCTATCTGGGATCGGGGAGGAAGACCCGCTTTTCCGCGAGGT 1138
 DB 2563 GGGATCACTTCAAACTTACCTTC-----AGCAGGTGCTTGTAGTTCCGCTC 2515
 QY 1139 TGAAGCGCATGGAAGAGTTTGGCAGATGACTGCTGCTGATTAAGCTTGAAGCA 1198
 DB 2514 ATCTTGAAAAATATATGTTCTTCTGTATCAATCTTCGGGATGCGACTTTGAAAAA 2455
 QY 1199 ACGCAGTTTACATATGATGTTGGGGAAGGTGGCCATGCAAGCCCTTAACGGTGA 1258
 DB 2454 GTCATGCCGTTTCAATATGATCTGGGTATCTTGAAGACATTTGAACACCATTAAGAAAGT 2395
 QY 1259 GTTCTGTTACAGCACTCGGATACAGTTGTCGCGGCTTTTCCGACACAGTTCCGAT 1318
 DB 2394 AGTGAACAAGTGTGGCATGGAACAGTAATTTTCCAGTAGTGAATAAATTAAGGAT 2335
 QY 1319 GGTACGCCGGAAGGCATCAAGAACCCGAACAATACCGGACAGACCGGAATGCGTGC 1378
 DB 2334 AAGTTTTC-----GTAATGTGATCATCTTCACCCCTTCACATGACAGAAAAATTTGTGC 2280
 QY 1379 CCGGTGTCATTAATGAACAGGGGTGGGCTGGGATTAATGATCAAGCAGAGACGGGTA 1438
 DB 2279 CATTAACATCACATCTTAATTAACAACAAATTTGGGCAACTCAAGTAAAGTTCTTCT 2220
 QY 1439 TCCGTGCTGATGCCGAGAAATGACATGATACCCCGTGAATACCCGCGCGCGCG 1498
 DB 2219 CTTTACTATACGGTACCAAGCTTTTGTTCCTTTAGTAGGGTTA-----ATTGCGCGC 2165
 QY 1499 TTGGCGTATCATGCTCATAGCTGTTTCTGTGTGAATTTTATTCGCTCACAATTTCA 1558
 DB 2164 TTGGCGTATCATGCTCATAGCTGTTTCTGTGTGAATTTTATTCGCTCACAATTTCA 2105
 QY 1559 CACAACATAGACCGGAACATTAAGTGAAGCCGTGGGCTGCTTAATGAGGATTA 1618

Db 2104 CACAACATAGAGCCGGAACATAAAGTAAAGCCTGGGGTGCCTTAATGAGTAGCTAA 2045
Qy 1619 CTACACATTAATTTGGGTTGGCTTCACTCCCGCTTTCAAGTGGGAAACCTGCTGCCAG 1678
Db 2044 CTACACATTAATTTGGGTTGGCTTCACTCCCGCTTTCAAGTGGGAAACCTGCTGCCAG 1985
Qy 1679 CTGCATTAATTAATTCGGCCCAAGCCGGGGAGAGCGGGTTTGGGTAATTTGGGCGCTTCC 1738
Db 1984 CTGCATTAATTAATTCGGCCCAAGCCGGGGAGAGCGGGTTTGGGTAATTTGGGCGCTTCC 1925
Qy 1739 GCTTCCTCGCTCACTGACTCGCTGGCTGGCTGGCTGGCTGGCGGAGGATACGCT 1798
Db 1924 GCTTCCTCGCTCACTGACTCGCTGGCTGGCTGGCTGGCTGGCGGAGGATACGCT 1865
Qy 1799 CACTCAAAAGCGGTAATACGGTTATCCAGAAATCAAGGGATAACCGAGAAAGAAATG 1858
Db 1864 CACTCAAAAGCGGTAATACGGTTATCCAGAAATCAAGGGATAACCGAGAAAGAAATG 1805
Qy 1859 TGAGCAAAAGCGCAGCAAAAGCGCAGGAACCGTAATAAGCCGCTTGGCGCTTTTC 1918
Db 1804 TGAGCAAAAGCGCAGCAAAAGCGCAGGAACCGTAATAAGCCGCTTGGCGCTTTTC 1745
Qy 1919 CATAGGCTCCGCCCCCTGACGAGCATCAAAATTCAGCTCAAGTAGAGGTGGCA 1978
Db 1744 CATAGGCTCCGCCCCCTGACGAGCATCAAAATTCAGCTCAAGTAGAGGTGGCA 1665
Qy 1979 AACCGAGAGAGATTAAGATACAGAGCGTTCCCTCGAAGCGCTCCCTGGTGGCTCT 2038
Db 1684 AACCGAGAGAGATTAAGATACAGAGCGTTCCCTCGAAGCGCTCCCTGGTGGCTCT 1625
Qy 2039 CTTGTTCCGACCCCTGGCTTACCGGATACCTGCTCCCTTCTCCCTGGGAAAGCGTG 2098
Db 1624 CTTGTTCCGACCCCTGGCTTACCGGATACCTGCTCCCTTCTCCCTGGGAAAGCGTG 1565
Qy 2099 GCGCTTCTCATAGCTCAAGCTGTAAGTCTCAAGTGGTGGTGGTGGTGGTGGTGGTGG 2158
Db 1564 GCGCTTCTCATAGCTCAAGCTGTAAGTCTCAAGTGGTGGTGGTGGTGGTGGTGGTGG 1505
Qy 2159 CTGGGCTGTGTGACGAACCCCGCTGAGCCCAACCGCTGCGCTTATCCGTAACAT 2218
Db 1504 CTGGGCTGTGTGACGAACCCCGCTGAGCCCAACCGCTGCGCTTATCCGTAACAT 1445
Qy 2219 CCGCTTGGAGTCCACCCGCTGAAGACGACTTATGCGCACTGGGAGGCACTGGTAAC 2278
Db 1444 CCGCTTGGAGTCCACCCGCTGAAGACGACTTATGCGCACTGGGAGGCACTGGTAAC 1385
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGGCTTAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGGCTTAAC 1325
Qy 2339 TACGGCTACATAGAAAGACATATTTTGTATCTGCGCTGCGTGAAGCAGTTACCTTC 2398
Db 1324 TACGGCTACATAGAAAGACATATTTTGTATCTGCGCTGCGTGAAGCAGTTACCTTC 1265
Qy 2399 GGAAGAAAGAGTTGATCTTGTATCCGCAACAAACGCGCTGGTGAAGCGGTGGTTT 2458
Db 1264 GGAAGAAAGAGTTGATCTTGTATCCGCAACAAACGCGCTGGTGAAGCGGTGGTTT 1205
Qy 2459 TTTGTTTGGCAAGCAGATTAACGCGCAAAAGAAAGATCTCAAGAGATCTTTGATC 2518
Db 1204 TTTGTTTGGCAAGCAGATTAACGCGCAAAAGAAAGATCTCAAGAGATCTTTGATC 1145
Qy 2519 TTTTCTTACCGGGCTTACGCTCAGTGGAAAGAAAGATCTCAAGAGATCTTTGATC 2578
Db 1144 TTTTCTTACCGGGCTTACGCTCAGTGGAAAGAAAGATCTCAAGAGATCTTTGATC 1085
Qy 2579 AGATTATCAAAAGAGATCTTCACTTAATCTTTTAAATTAATTAATTAATTAATCA 2638
Db 1084 AGATTATCAAAAGAGATCTTCACTTAATCTTTTAAATTAATTAATTAATTAATCA 1025
Qy 2639 ATCTTAAGTATATAGTAAACCTTGTGACAGTTAACAGTCTTAATCAAGTAGGCA 2698
Db 1024 ATCTTAAGTATATAGTAAACCTTGTGACAGTTAACAGTCTTAATCAAGTAGGCA 965

Qy 2699 CCTATCTAGCGATCTGTATATTTGTTATCCATAGTAGTGGCTGATCCCGCGGTGAG 2758
Db 964 CCTATCTAGCGATCTGTATATTTGTTATCCATAGTAGTGGCTGATCCCGCGGTGAG 905
Qy 2759 ATTAATCAGATACGGAGGCTTACCATCTGGGCCCCAGTGTGCAATGATACCGGAGAC 2818
Db 904 ATTAATCAGATACGGAGGCTTACCATCTGGGCCCCAGTGTGCAATGATACCGGAGAC 845
Qy 2819 CCAGGCTACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAAAGGCGAGGCG 2878
Db 844 CCAGGCTACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAAAGGCGAGGCG 785
Qy 2879 AGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCT 2938
Db 784 AGAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCT 725
Qy 2939 AGAGTAAGTATTCGCACTTAATAGTTTGGCAAGTTGTCATGCTACAGGATC 2998
Db 724 AGAGTAAGTATTCGCACTTAATAGTTTGGCAAGTTGTCATGCTACAGGATC 665
Qy 2999 GTGGTGTACGCTCGTCTGTTGTTATGGCTTATTCAGCTCCGTTCCAGATCAAGG 3058
Db 664 GTGGTGTACGCTCGTCTGTTGTTATGGCTTATTCAGCTCCGTTCCAGATCAAGG 605
Qy 3059 CGAGTTACATGATCCCCATGTTGTGCAAAAGCGGTAGCTCTTCGCTCCGATC 3118
Db 604 CGAGTTACATGATCCCCATGTTGTGCAAAAGCGGTAGCTCTTCGCTCCGATC 545
Qy 3119 GTTGTCAAGATTAAGTTGGCGCAGGTATTCATGTTATGTCAGACATGCAATAT 3178
Db 544 GTTGTCAAGATTAAGTTGGCGCAGGTATTCATGTTATGTCAGACATGCAATAT 485
Qy 3179 TCTCTTCTGTGACGCACTCCGTAAGTCTTGTGACTGTGAGTACTCAAG 3238
Db 484 TCTCTTCTGTGACGCACTCCGTAAGTCTTGTGACTGTGAGTACTCAAG 425
Qy 3239 TCATTCTGAAGATAGTGTATGCGCGACGAGTGTCTTGGCGCGCTCAATACGGAT 3298
Db 424 TCATTCTGAAGATAGTGTATGCGCGACGAGTGTCTTGGCGCGCTCAATACGGAT 365
Qy 3299 AATACGCGCGCATAGCAGAACTTTAAAGTGCTCATTTGAAAGCTTCTGGGG 3358
Db 364 AATACGCGCGCATAGCAGAACTTTAAAGTGCTCATTTGAAAGCTTCTGGGG 305
Qy 3359 CGAAAACCTGCAAGATCTTACCGCTGTTGAGATCCAGTTGAGATACCACTCGTGA 3418
Db 304 CGAAAACCTGCAAGATCTTACCGCTGTTGAGATCCAGTTGAGATACCACTCGTGA 245
Qy 3419 CCCAATGATCTTACGCACTTTTACTTCAACAGCGTTCTGGGTGAGCAAAACAGG 3478
Db 244 CCCAATGATCTTACGCACTTTTACTTCAACAGCGTTCTGGGTGAGCAAAACAGG 185
Qy 3479 AGGCAAAATGCGCAAAAGGAAATAGGCGGCAAGGAAAGTTGATATCTCATATC 3538
Db 184 AGGCAAAATGCGCAAAAGGAAATAGGCGGCAAGGAAAGTTGATATCTCATATC 125
Qy 3539 TTCTTTTCAATTAATTAAGATTTATCAGGTTATTTGTCTCATGAGCGGATACAT 65
Db 124 TTCTTTTCAATTAATTAAGATTTATCAGGTTATTTGTCTCATGAGCGGATACAT 65
Qy 3599 TTTGAATGATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAATTTCCGAAAAGTG 3658
Db 64 TTTGAATGATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAATTTCCGAAAAGTG 5
Qy 3659 CCAC 3662
Db 4 CCAC 1

RESULT 10
AD126329/c
ID AD126329 standard; DNA; 3637 BP.

XX AD126329;
AC
XX 22-APR-2004 (first entry)
XX
XX Novel endonuclease Res I-related plasmid pBSC3BFP.
XX
XX endonuclease; molecular biology; plant propagation; phenotypic trait;
XX herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;
XX osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
XX bacterial; viral; enzyme production; secondary metabolite;
XX male sterility; female sterility; dwarfness; early maturity; Res I;
XX pBSC3BFP; de.
XX
XX Unidentified.
XX
XX US2003148315-A1.
XX
XX 07-AUG-2003.
XX
XX 01-AUG-2002; 2002US-00211079.
XX
XX 01-FEB-2002; 2002US-0353722P.
XX
XX 14-MAR-2002; 2002US-00098155.
XX
XX (PADG/) PADGETT H S.
XX (VAEM/) VAEMHONGS A A.
XX
XX Padgett HS, Vaemhongs AA;
XX
XX WPI, 2003-897548/82.
XX
XX
XX New nucleic acid molecule encoding endonucleases, useful in molecular
XX biology, specifically to generating populations of related nucleic acid
XX molecules, and in plant propagation with useful phenotypic traits.
XX
XX Example 14; Fig 5; 46pp; English.
XX
XX This invention relates to a novel endonuclease (Res I) nucleic acid
XX molecule which comprises a fully defined sequence of 899 bp given in the
XX specification. The methods and compositions of the present invention are
XX useful in molecular biology, and more specifically to generating
XX populations of related nucleic acid molecules. They may also be used in
XX plant propagation with useful phenotypic traits, such as improved
XX tolerance to herbicides, improved tolerance to extremes of heat or cold,
XX drought, salinity or osmotic stress, improved resistance to pests
XX (insects, nematodes or arachnids) or diseases (fungal, bacterial or
XX viral), production of enzymes or secondary metabolites, male or female
XX sterility, dwarfness and early maturity. The present sequence is that of
XX a plasmid which is related to the invention.
XX
XX Sequence 3637 BP; 942 A; 900 C; 860 G; 935 T; 0 U; 0 Other;
SQ
Query Match 72.3%; Score 2647.8; DB 10; Length 3637;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3054; Conservative 0; Mismatches 582; Indels 27; Gaps 3;
QY 1 CTAATGTGAAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
DB 3637 CTAATGTGAAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAATTCGCAAAATCCCTTAATTAATTAAGAAATAGACGA 120
DB 3577 ATTTTAAACCAATAGCGCGAATTCGCAAAATCCCTTAATTAATTAAGAAATAGACGA 3518
QY 121 GATAGGGTGAAGTGTGTTCCAGTTTGAACAAGAGTCCATTAATTAAGAAAGTGAAGT 180
DB 3517 GATAGGGTGAAGTGTGTTCCAGTTTGAACAAGAGTCCATTAATTAAGAAAGTGAAGT 3458
QY 181 CAAGCTCAAGGGCGAAGAAACCGTCTATCAAGGGCGAGTCCCACTAAGCAATCAAC 240
DB 3457 CAAGCTCAAGGGCGAAGAAACCGTCTATCAAGGGCGAGTCCCACTAAGCAATCAAC 3398

QY 241 CTAATCAAGTTTGTGGGGTGAAGTGCCTTAAGCACTAAATGGAACCTTAAGGAG 300
DB 3397 CTAATCAAGTTTGTGGGGTGAAGTGCCTTAAGCACTAAATGGAACCTTAAGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAAGCCCGCAACGTGCGAAGAAAGAAAGAA 360
DB 3337 CCCCCGATTTAGAGCTTGAACGGGGAAGCCCGCAACGTGCGAAGAAAGAAAGAA 3278
QY 361 AGCGAAGAGAGCGGCGCTAGGAGCGCTGGCAAGTGAAGCGGTCAAGCTGCGTAAAC 420
DB 3277 AGCGAAGAGAGCGGCGCTAGGAGCGCTGGCAAGTGAAGCGGTCAAGCTGCGTAAAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTAAGCGGCGCGCTCCATTCGCAATCAGCTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTAAGCGGCGCGCTCCATTCGCAATCAGCTGCG 3158
QY 481 CAACTGTGGGAAGGCGATGCTGCGGCGCTCTTCCGTTATTAAGCCAGTGGGAAAG 540
DB 3157 CAACTGTGGGAAGGCGATGCTGCGGCGCTCTTCCGTTATTAAGCCAGTGGGAAAG 3098
QY 541 GGAATGTGCTGAAGGCGATTAAGTTGGGTAAACCCAGGTTTCCAGTCAAGCTGG 600
DB 3097 GGAATGTGCTGAAGGCGATTAAGTTGGGTAAACCCAGGTTTCCAGTCAAGCTGG 3038
QY 601 TAAACGACGCGCGTGAAGCGCGCTGTTCAATTCAGTTTGAACCCGCGAGGACG 660
DB 3037 TAAACGACGCGCGTGAAGCGCGCTGTTCAATTCAGTTTGAACCCGCGAGGACG 2978
QY 661 GCAGACTGCGCGTCAAAATGTGTTTACAGCGTATGAGAGATGAAGTGTGACAC 720
DB 2977 CCGGCGTGGCGCGCGCTGTAAGTGAAGTATCCCGCGCTGAGAGATTTTATTTGT 2918
QY 721 GCTGCAAGAACGACGAGTGAATTAACCTTAAGAAAGATTAATTAAGTGAAGTGA 780
DB 2917 ATAGTTATTCATCCATCCATGTTGTAATCCAGACGAGTTAACTCAAGAAAGACATGT 2858
QY 781 AGATTAATC-ATGCGTAAATTAAGCAAGTGGATGTTGAATTAAGTCACTAATAGGCGAA 839
DB 2857 GGTACAGCTTTTCTTGGGATCTTTGGAAGGAGATGTTGTCAGACGTTAAGTGTGT 2798
QY 840 TTGGGTACCGGCGCGCGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 899
DB 2797 CTGGTAAAGAGACAGGCGCATTCGCAATTAAGTGAAGTGAAGTGAAGTGAAGT 2738
QY 900 CAGCCCGGAGGATCAGTGAAGTGAAGGCGCGCACCGCGTGAAGTGAAGTGAAGT 959
DB 2737 GAAAGGATTCATCTTCAATGTGCGGAAATTTGAAGTGAAGTGAAGTGAAGT 2678
QY 960 TTCCCTTAAGAGAGGTTAATTAATCCATGCTCAATTTAGCAGACATATCTTTCTA 1019
DB 2677 GTTGTCTGCGCGTATGATATATATGTTGAGTTAAAGTGTATCTGAGTTGTGTCGA 2618
QY 1020 GGGTTAATCTAGCTGATCAGATCATATGTCGGGCTTTTCCGGCTCAGTCATGCGC 1079
DB 2617 GAAAGTTCATCTTCTTTAAATCAATTAATCTTTAATCTGATC-----GATTAACAG 2563
QY 1080 CCAAGCTGGCGCTATCTGGGATCGGAGAGGAGAAAGACCCGTGCTTTTCCCGAGGTT 1139
DB 1140 GAAAGCGCATGGAAGAGTTTCCGAGAGATGATGCTGCTGATTAAGCTTGAAGCAAAA 1199
QY 2562 TATAGTCGTTCTGTCATATACCTTGGGCGATGCGCTTGAAGAAAGTCATGCGGTTT 2443
DB 1200 CGCAGCTTTACATGATGATTTCCGGAAGTGTGCGCAATGACGCTTTAAGCGTGAAGT 1259
QY 2442 CATATGATTCGGAATTAAGAAAGATTAAGCAATTAAGCAATTAAGCAATTAAGCA 2383
DB 1260 TTGCTTCAAGGCACTGGGATTAACGATTTGCTGCGGCTTTTCCGAGACAGATTCG 1319
QY 2382 TGCCATAGGAACAGGTGTTTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2323
DB 1320 GTCAAGCCGAAGCGCATTCAGCAACCGGAACATTCGCGGCAACGCGGAACTGCGGTC 1379

Db 2322 TGTAGATACCTTCACTCTTCCACTGACGAAATTTGTGCCATTAATACATC 2263
Qy 1380 GGTGTCAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTACAGAGCGGTTAT 1439
Db 2262 TTAATTCACAAAGAT-----TGGACAACTCCAGTGAATAATTC 2224
Qy 1440 CCTGCTGGATGCGCAGAAATGACATGATACCCCGTAGATTACCCGCGGCGGCT 1499
Db 2223 TTCTCTTTATCTCATCGGTACCCAGCTTTGTGTGCTTTAGTAGAGGTTAATTGGCGGCT 2164
Qy 1500 TGGCGTATCATGTGATAGCTGTTTCTGTGTAATTTGTATCCGCTACAAATTCAC 1559
Db 2163 TGGGTATCATGTGATAGCTGTTTCTGTGTAATTTGTATCCGCTACAAATTCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCTAAAGTAAAGCTGGGGTGTCTAATAGTAGAGTAAC 1619
Db 2103 ACACATACGAGCCGGAAGCTAAAGTAAAGCTGGGGTGTCTAATAGTAGAGTAAC 2044
Qy 1620 TCAATTAATTCGCTGCTGCTCACTGCGCTTTCCAGTCGGGAAACCTGTGTCAGC 1679
Db 2043 TCAATTAATTCGCTGCTGCTCACTGCGCTTTCCAGTCGGGAAACCTGTGTCAGC 1984
Qy 1680 TGCATTATGAATCGGCAACGCGCGGAGAGCGGTTTGCATATTGGGCGCTTTCCG 1739
Db 1983 TGCATTATGAATCGGCAACGCGCGGAGAGCGGTTTGCATATTGGGCGCTTTCCG 1924
Qy 1740 CTTCCTGCTCACTGACTGCTGCGCTGCGTTCGCTGCGTGGGAGACGCTATGACTC 1799
Db 1923 CTTCCTGCTCACTGACTGCTGCGTTCGCTGCGTTCGCTGCGGAGACGCTATGACTC 1864
Qy 1800 ACTCAAAAGCGGTAAATACGTTATCAAGAAATCGAGGATTAACGACAGAAAGAACTGT 1859
Db 1863 ACTCAAAAGCGGTAAATACGTTATCAAGAAATCGAGGATTAACGACAGAAAGAACTGT 1804
Qy 1860 GAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGACGCGTTCGCTGCTTTTCC 1919
Db 1803 GAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGACGCGTTCGCTGCTTTTCC 1744
Qy 1920 ATAGCTCCGCCCCCTTACAGAGCTTCAGAAATTCAGCGTCAAGTCAGAGGTGGGAA 1979
Db 1743 ATAGCTCCGCCCCCTTACAGAGCTTCAGAAATTCAGCGTCAAGTCAGAGGTGGGAA 1684
Qy 1980 ACCGACAGACCTAATAAGATACAGAGGTTTCCCGTGAAGTCCCTGTGCGCTTC 2039
Db 1683 ACCGACAGACCTAATAAGATACAGAGGTTTCCCGTGAAGTCCCTGTGCGCTTC 1624
Qy 2040 CTGTTCCGACCCCTGCGTTACCGGATACCTGTCCGCTTTCTCCCTGGGAAAGCTG 2099
Db 1623 CTGTTCCGACCCCTGCGTTACCGGATACCTGTCCGCTTTCTCCCTGGGAAAGCTG 1564
Qy 2100 CGCTTTCTCATAGCTACGCTGTAGATCTCAGTTCCGATGAGGTCCGTTCCGCTCAAGC 2159
Db 1563 CGCTTTCTCATAGCTACGCTGTAGATCTCAGTTCCGATGAGGTCCGTTCCGCTCAAGC 1504
Qy 2160 TGGGCTGTGTCAGAAACCCCGCTTCAAGCCGACCGCTGCGCTTATCCGTAATCTATC 2219
Db 1503 TGGGCTGTGTCAGAAACCCCGCTTCAAGCCGACCGCTGCGCTTATCCGTAATCTATC 1444
Qy 2220 GTCTTGATCCAAACCGGTAAAGACAGACTTTCGCACTGGCAGACGACTGTGTAACA 2279
Db 1443 GTCTTGATCCAAACCGGTAAAGACAGACTTTCGCACTGGCAGACGACTGTGTAACA 1384
Qy 2280 GATTTAGCAGAGGAGGTATGTAAGCGGTGTCAAGAGTTCTTGAAGTGGTGGCTAACT 2339
Db 1383 GATTTAGCAGAGGAGGTATGTAAGCGGTGTCAAGAGTTCTTGAAGTGGTGGCTAACT 1324
Qy 2340 ACGGCTACACTAAGAGACAGTATTGGATCTGGGCTCTGCTGAAGCACTTTCG 2399
Db 1323 ACGGCTACACTAAGAGACAGTATTGGATCTGGGCTCTGCTGAAGCACTTTCG 1264
Qy 2400 GAAAAAGATTGTGCTTGTATCCGCAAAACAACCAACGCTGTAGCGGTGTTT 2459

Db 1263 GAAAAAGATTGTAGCTTGTATCCGGCAAAACAACCCGCTGAGCGGTGTTT 1204
Qy 2460 TTGTTGCAACACACATTAATACGCGCAAAAAAAGATTCGAAGATTCCTTGTATCT 2519
Db 1203 TTGTTGCAACACACATTAATACGCGCAAAAAAAGATTCGAAGATTCCTTGTATCT 1144
Qy 2520 TTTTCAGGGGTCTGACGCTCAGTGGAAAGAAATCAGGTTAAGGATTTTGTATCA 2579
Db 1143 TTTTCAGGGGTCTGACGCTCAGTGGAAAGAAATCAGGTTAAGGATTTTGTATCA 1084
Qy 2580 GATTATCAAAAAGATCTTCACTTAATCTTTTAAATTAATTAATCA 2639
Db 1083 GATTATCAAAAAGATCTTCACTTAATCTTTTAAATTAATTAATCA 1024
Qy 2640 TCTAAGATATATATGATTAACCTTGGTCTGACAGTTACCAATGCTAATCAGTAGC 2699
Db 1023 TCTAAGATATATATGATTAACCTTGGTCTGACAGTTACCAATGCTAATCAGTAGC 964
Qy 2700 CTATCTCAGGATCTGTCTAATTTGTTTCAATCCATAGTTGCTGACTCCCGCTGTAGA 2759
Db 963 CTATCTCAGGATCTGTCTAATTTGTTTCAATCCATAGTTGCTGACTCCCGCTGTAGA 904
Qy 2760 TAACTACGATACGGAGGAGGCTTAACATCTGGCCCAAGTCTGCAATATCCGAGACC 2819
Db 903 TAACTACGATACGGAGGAGGCTTAACATCTGGCCCAAGTCTGCAATATCCGAGACC 844
Qy 2820 CAGGCTCACCGGCTCCATTTATTCAGATTAACCAAGCCAGCCGAGGCGGAGCCA 2879
Db 843 CAGGCTCACCGGCTCCATTTATTCAGATTAACCAAGCCAGCCGAGGCGGAGCCA 784
Qy 2880 GAAATGTCCTGCAACTTTATCCGCTCCATCCAGTCTAATTTGTGCGGAAAGCTA 2939
Db 783 GAAATGTCCTGCAACTTTATCCGCTCCATCCAGTCTAATTTGTGCGGAAAGCTA 724
Qy 2940 GAGTAAGATTCGCGGATTAATGTTTGGCAACGTTGTTGCAATGCTACAGGATCG 2999
Db 723 GAGTAAGATTCGCGGATTAATGTTTGGCAACGTTGTTGCAATGCTACAGGATCG 664
Qy 3000 TGTGTCAACGCTCGTCTTTTGTATGAGCTTCATTGAGCTCCGTTCCCAAGATCAAGC 3059
Db 663 TGTGTCAACGCTCGTCTTTTGTATGAGCTTCATTGAGCTCCGTTCCCAAGATCAAGC 604
Qy 3060 GAGTTACATATCCCGCATATGTTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCG 3119
Db 603 GAGTTACATATCCCGCATATGTTGTGCAAAAAGCGTTAGCTCTTCGCTCCGATCG 544
Qy 3120 TTGTCAAGATTAAGTTGGCGCAGTGTATCTCATGAGTTATGGCAGACCTGATTAAT 3179
Db 543 TTGTCAAGATTAAGTTGGCGCAGTGTATCTCATGAGTTATGGCAGACCTGATTAAT 484
Qy 3180 CTCTACGTGATGCAATCCGTAAGTGTCTTGTGATAGTGTGATCAACCAAGT 3239
Db 483 CTCTACGTGATGCAATCCGTAAGTGTCTTGTGATAGTGTGATCAACCAAGT 424
Qy 3240 CATTTGAAATAGTGTATGCGGACGAGTGTGCTTGGCGCGGCTCAATACGGATA 3299
Db 423 CATTTGAAATAGTGTATGCGGACGAGTGTGCTTGGCGCGGCTCAATACGGATA 364
Qy 3300 ATACCGCGCACAATGCGAACTTTAAAGTGCATCATTTGAAAGAGTTCTTCGGGG 3359
Db 363 ATACCGCGCACAATGCGAACTTTAAAGTGCATCATTTGAAAGAGTTCTTCGGGG 304
Qy 3360 GAAAACTCTCAAGATCTTACCGCTGTGAGTCAATTTGATGATTAACCATCTGTCAC 3419
Db 303 GAAAACTCTCAAGATCTTACCGCTGTGAGTCAATTTGATGATTAACCATCTGTCAC 244
Qy 3420 CCAATGATCTTCAGCATCTTTTATCTTTACAGCGTTTCTGGGTGAGCAAAACAGGAA 3479
Db 243 CCAATGATCTTCAGCATCTTTTATCTTTACAGCGTTTCTGGGTGAGCAAAACAGGAA 184
Qy 3480 GGCAAAATGCGCAAAAAGGAAATAGGGCGCACGAAATGTTGAATATCATATCT 3539
Db 183 GGCAAAATGCGCAAAAAGGAAATAGGGCGCACGAAATGTTGAATATCATATCT 124

QY 3540 TCCTTTCAATATATGAGCATTTATCAGGTTATTCATGAGCGATATCATAT 3599
DB 123 TCCCTTTTCAATATATGAGCATTTATCAGGTTATTCATGAGCGATATCATAT 64
QY 3600 TTGAATGATTTAGAAAATAAACAATAGGGGTTCCCGCAGCATTTCCCGAAAAGTGC 3659
DB 63 TTGAATGATTTAGAAAATAAACAATAGGGGTTCCCGCAGCATTTCCCGAAAAGTGC 4
QY 3660 CAC 3662
DB 3 CAC 1
RESULT 11
AAQ13578
ID AAQ13578 standard; DNA; 3681 BP.
AC AAQ13578;
XX 02-DEC-1991 (first entry)
DT 02-DEC-1991 (first entry)
XX
DE Plasmid pKSEB5.
XX recombinant expression plasmid; phage phi X174; protein E;
KW bacterial ghost; 88.
XX
OS Synthetic.
XX MO9113155-A.
PN
XX 05-SEP-1991.
PD
XX 24-FEB-1990; 90DE-04005874.
PF 24-FEB-1990; 90DE-04005874.
XX
PR 24-FEB-1990; 90DE-04005874.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
P1 Lubitz W, Szostak MP;
DR WPI; 1991-281471/38.
XX
PT Carrier recombinant protein for vaccines against HIV, etc. - obtd. by
FT expressing fusion protein gene in gram negative bacteria and gene that
XT encodes for lytic membrane protein.
PS
XX Example 1; Page 33-34; 45pp; German.
CC Plasmid pKSEB5 comprises three multiple cloning sites (mccl-3), the
CC ampicillin resistance gene, regions of the lac operon and a partial
CC phage phi X174 E' target sequence. This plasmid can be used as a carrier
CC vector according to the invention for the insertion of sequences encoding
CC viral antigenic peptides, e.g. the HIV gp41 sequence was isolated as a
CC HindIII/PvuII fragment from plasmid pHF14. The fragment also included a
CC linker sequence and the last 45 codons of gp120. The fragment was ligated
CC to pKSEB5 which had been cleaved with AccI and end-filled. Alternatively,
CC a new carrier vector, denoted pMTV1 (see AAQ13576), can be constructed
CC from pKSEB5 which has been partially digested with SmaI. It is ligated to
CC a SmaI-DraIII fragment from pMT1 (see AAQ13577). Foreign sequences are
CC cloned into mcs2
XX
SQ Sequence 3681 BP; 949 A; 917 C; 932 G; 883 T; 0 U; 0 Other;
Query Match 72.0%; Score 2635.8; DB 2; Length 3681;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3073; Conservative 0; Mismatches 582; Indels 28; Gaps 6;
QY 3 AAATGTTAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 62
DB 1 AAATGTTAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGCTCAT 60
QY 63 TTTTAAACAATAGCGCAAAATCCCTTAATCAAAAAGATAGACCGAGA 122

DB 61 TTTTAAACAATAGCGCAAAATCCCTTAATCAAAAAGATAGACCGAGA 120
QY 123 TAGGTTGAGTGTGTTTCACTTGGAAACAAGTCCATTTAAAGACGTGACCTCA 182
DB 121 TAGGTTGAGTGTGTTTCACTTGGAAACAAGTCCATTTAAAGACGTGACCTCA 180
QY 183 ACGTCAAAAGGCGAAAACCGTCTATCAAGGCGCATGCGCCACTAGTAAACATCACCT 242
DB 181 ACGTCAAAAGGCGAAAACCGTCTATCAAGGCGCATGCGCCACTAGTAAACATCACCT 240
QY 243 AATCAAGTTTGTGGGTCGAGTGCCTTAAGCACTTAATCGGAACCTTAAGGAGACC 302
DB 241 AATCAAGTTTGTGGGTCGAGTGCCTTAAGCACTTAATCGGAACCTTAAGGAGACC 300
QY 303 CCCGATTTAGAGCTTGAACGGGAAAGCCGGCAAGTGGCGAAGAAAGGAAAGAA 362
DB 301 CCCGATTTAGAGCTTGAACGGGAAAGCCGGCAAGTGGCGAAGAAAGGAAAGAA 360
QY 363 CGAAAGAGCGGGCGCTAGAGGCGCTGGCAAGTGAAGCGGTGACGCTGGCGTAAACACCA 422
DB 361 CGAAAGAGCGGGCGCTAGAGGCGCTGGCAAGTGAAGCGGTGACGCTGGCGTAAACACCA 420
QY 423 CACCGCGCGGCTTAATAGCGCGCTTACAGGCGCGCTCCATTCGCAATTCAGGCTGCGCA 482
DB 421 CACCGCGCGGCTTAATAGCGCGCTTACAGGCGCGCTCCATTCGCAATTCAGGCTGCGCA 480
QY 483 ACTGTTGGAAAGGCGATCGGTGCGGCGCTTTCGCTATTAAGCGAGCTGCGAAAGGGG 542
DB 481 ACTGTTGGAAAGGCGATCGGTGCGGCGCTTTCGCTATTAAGCGAGCTGCGAAAGGGG 540
QY 543 GATGTGCTGCAAGCGCATTTAAGTGGTAAACGCCAGGGTTTCCCAAGTCAAGAGTGTGA 602
DB 541 GATGTGCTGCAAGCGCATTTAAGTGGTAAACGCCAGGGTTTCCCAAGTCAAGAGTGTGA 600
QY 603 AAACGAGCGGCGAGTGAAGCGCGCTCGTTCAATTCAGTTTGAACCGGTGAGGAGCGGCG 662
DB 601 AAACGAGCGGCGAGTGAATGTTATTAAGCATCTATTAAGGCGAAATGGAGCTCCACCGG 660
QY 663 AGACTCGCGGTGCAATGTGTTTACAGCGGTGATGAGCAGATGAAGATGCTGACAGCG 722
DB 661 GTGGCGGCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 723 TGCAGAAACGACAGCTGATTAACCTTAAGAAATTAATCATATTGTGACGTAAAG 782
DB 721 AGCCAGTATATACACTCGGCTATGCTAGTACGTAGGAGTCACTGCTGCGCCACACCC 780
QY 783 AATATCAGGCTAAATATGAAGCATGAGATCTGTAATACAGCTCATATAGGCGGAATG 842
DB 781 GCCAACACCCGCTGACGCGGCTTGAAGCGGCTTGTCTGCTCCCGCATCCCTTACAGACA 840
QY 843 GGTACCGGCGCCCGCTCGAGTGCAGCGATACGA-----TAAGCTTGATATGAAATTC 896
DB 841 AGCTGTACGCTCTCGGAGCTGATGTCAGAGGTTTTCACCGTCATACCGAAAGC 900
QY 897 CTGACGCCGCGGAGTCACTAGTTCTAAGCGGCGCCACCGCGGTGAGCTCAAGTT 956
DB 901 CGCAGGCAAGTAAAGTGCAGTCTTGTGAGCAATTCGCTTAAAGCAATTTGCTG 960
QY 957 TTGTTCCTTTAGAGAGGTTAATTAATCCCAT-----GGCTCAATTTAAGCAGAC 1009
DB 961 TAAAGTGTACCTGTGCGATACCGCTTCCAGTAGGACAGAAAGCAATGATGTAA 1020
QY 1010 TATCTTTTAGGTTAATCTAGCTCATAGATCATATCGTCCGGTCTTTTTCGCGCT 1069
DB 1021 TTTTGAGAGAAAGATCGAGAGAGATCATATTAAGAGTTGAACCTTCTTTGTGCT 1080
QY 1070 CAGTCAATGCCCAAGCTGGGCTATCTGGGCTATGGGAGAGAGAACCGCTGCTTTTC 1129
DB 1081 TCAGCATAGGTAAATCCATATGTTGAATGGCCCTAGAGATCCCGCAAGCTTCCATAGCC 1140
QY 1130 CCGGAGGTTGAACGCGCATAGAAAGATTTGCCGAGATGATGCTGCTGATTAAGCT 1189

Db 1141 TCGAGGTGACCTTAGAGATCCCGACGCTCGAGCCCATTAATATGTTTTCCGTAAT 1200
QY 1190 TTAGGAAAAAGCAGCTTAAACATGATGATTCGGGAAAGTGTGGCAATGACCGCTTTAA 1249
Db 1201 TCAAGGCTTCCATGATGAGACAGGCCGTTT--GAAAGTTGACGGGATGAACATAAAG 1258
QY 1250 CGGTAACTGTTTCGTCAGGCGACCTGGGATACAGATTCCGCGGCTTTTCCGAGACA 1309
Db 1259 CAATGACGGGACGAAATAAATCAACAGG---AGAGAGAAAGCAGAGGATATCCACAAGT 1315
QY 1310 GTTCCGATGTGACGCCGAAAGCGCATCAGCAACCCGAAATATCCGGGAGAG----- 1363
Db 1316 CCAGGTACCATTAACCAAGCTCAACGACGAGGACGAGACGAGCCGATGATGCAA 1375
QY 1364 CCGGAATGCGCGTCCGGTGTGAGATTATGACAGGGGTGCGGCGCTGGGATATTCGT 1423
Db 1376 TCCAAACTTTGTTACTGTCAAGAAATCGAAATCATCTTCGGTTAAATCAAAACGGCAG 1435
QY 1424 CAGGAGAGAGGGTAT---CTTGGCTGATGCCGAGAAATGACATGGATATCCCGTG 1479
Db 1436 AAGCTGAATGAGAAATCGACCTCGAGGGGGGCCGGTACCAGCTTTTGTCCCTTTA 1495
QY 1480 AGTTACCCGGCGGGCGGCTTGGCGCTTAATCATGTGTATAGCTGTTTCTGTGTGAATTG 1539
Db 1496 GTAGAGGTTAATTCAGACTTGCGTGAATCATGTGTATAGCTGTTTCTGTGTGAATTG 1555
QY 1540 TTATCCGCTCAAAATTCACACAAACATACAGACCGGAGCATAAAGTGTAAAGCTTGAG 1599
Db 1556 TTATCCGCTCAAAATTCACACAAACATACAGACCGGAGCATAAAGTGTAAAGCTTGAG 1615
QY 1600 TGCTTAATAGTAGTACTCACTCACTTAATTCGTTGCGCTCACTGCGCGCTTTCAGTTC 1659
Db 1616 TGCTTAATAGTAGTACTCACTCACTTAATTCGTTGCGCTCACTGCGCGCTTTCAGTTC 1675
QY 1660 GGGAAACCTGTGTGCAGCTGCACTTAATGAAATCGGACCAAGCGGGGAGAAAGCGGTTT 1719
Db 1676 GGGAAACCTGTGTGCAGCTGCACTTAATGAAATCGGACCAAGCGGGGAGAAAGCGGTTT 1735
QY 1720 GGGTATGGGCGCTCTTCGCGTCTCTGCTCACTGACTGCTGCGTCTGCTGCTGCTGCT 1779
Db 1736 GGGTATGGGCGCTCTTCGCGTCTCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCT 1795
QY 1780 GCGGAGACGGTATCACTCACTAAAGCGGTAATCGGTTATCCACAAATTCAGGGCA 1839
Db 1796 GCGGAGACGGTATCACTCACTAAAGCGGTAATCGGTTATCCACAAATTCAGGGCA 1855
QY 1840 TAAAGCAGGAGAAACATGTAGCAAAAGCCAGCAAAAGCCAGAAACGTTAAAGGC 1899
Db 1856 TAAAGCAGGAGAAACATGTAGCAAAAGCCAGCAAAAGCCAGAAACGTTAAAGGC 1915
QY 1900 CGGTTGCTGCGCTTTTTCATAGGCTCCGCCCCCGGAGACATCACAAAATTCAGAGC 1959
Db 1916 CGGTTGCTGCGCTTTTTCATAGGCTCCGCCCCCGGAGACATCACAAAATTCAGAGC 1975
QY 1960 CTAAAGTCAGAGTGTGGGAAACCCGACAGGATATTAAGATACAGGCGTTTCCCGCTG 2019
Db 1976 CTAAAGTCAGAGTGTGGGAAACCCGACAGGATATTAAGATACAGGCGTTTCCCGCTG 2035
QY 2020 AAGCTCCCTGTGCGCTCTCTGTTCCGACCCGCGGCTTAACCGGATACCTGTCCGCTT 2079
Db 2036 AAGCTCCCTGTGCGCTCTCTGTTCCGACCCGCGGCTTAACCGGATACCTGTCCGCTT 2095
QY 2080 TCTCCCTTGGGAAAGGTGGCGCTTTCTCATAGCTCAAGCTTAGAGTATCTCAGTTGGGT 2139
Db 2096 TCTCCCTTGGGAAAGGTGGCGCTTTCTCATAGCTCAAGCTTAGAGTATCTCAGTTGGGT 2155
QY 2140 GTAGTGTGTTGCTCAAGCTGGGCTGTGTGACAAACCCCGCTTCAAGCCGACCGCTG 2199
Db 2156 GTAGTGTGTTGCTCAAGCTGGGCTGTGTGACAAACCCCGCTTCAAGCCGACCGCTG 2215
QY 2200 CGCTTATCCGGTAATCTGTTGAGTCCAAACCCGGTAAGACAGCATTAATGCGCACT 2259
Db 2216 CGCTTATCCGGTAATCTGTTGAGTCCAAACCCGGTAAGACAGCATTAATGCGCACT 2275

QY 2260 GGCAGACCACTGGTAAACAGATTACACAGCGGATATGTAAGCGCTGTCAACAGATT 2319
Db 2276 GGCAGACCACTGGTAAACAGATTACACAGCGGATATGTAAGCGGCTGTCAACAGATT 2335
QY 2320 CTGTAAGTGTGGCTTAACCTACCGCTACACTAGAAAGACAGTATTTGTATCTGCGCTC 2379
Db 2336 CTGTAAGTGTGGCTTAACCTACCGCTACACTAGAAAGACAGTATTTGTATCTGCGCTC 2395
QY 2380 GCTAAGCCAGTTAATCTTGGGAAAGAGTGTACTCTTGAATCCGGCAACAAACGAC 2439
Db 2396 GCTAAGCCAGTTAATCTTGGGAAAGAGTGTACTCTTGAATCCGGCAACAAACGAC 2455
QY 2440 CGCTGTAGAGGCTGTGTTTTTTTGTGCAAGCAGCATTAACGCGCAAAAGAGATC 2499
Db 2456 CGCTGTAGAGGCTGTGTTTTTTTGTGCAAGCAGCATTAACGCGCAAAAGAGATC 2515
QY 2500 TCAAGAAATCTTTTATCTTTTCTAAGGGGTCTGACGCTCAGTGAACGAAATCTCAG 2559
Db 2516 TCAAGAAATCTTTTATCTTTTCTAAGGGGTCTGACGCTCAGTGAACGAAATCTCAG 2575
QY 2560 TTAAAGGATTTGTGATGATATTAACAAAGAGATCTTCACTAGATCTTTTAAATTA 2619
Db 2576 TTAAAGGATTTGTGATGATATTAACAAAGAGATCTTCACTAGATCTTTTAAATTA 2635
QY 2620 AAAATGAATTTTAAATCAATCTAAGTATATGAGTAAACCTTGTCTGACATTAACA 2679
Db 2636 AAAATGAATTTTAAATCAATCTAAGTATATGAGTAAACCTTGTCTGACATTAACA 2695
QY 2680 ATGCTTAATCAATGAGGACCTATCTCAGCATCTGTCTATTTGCTTATCATCATAGTTG 2739
Db 2696 ATGCTTAATCAATGAGGACCTATCTCAGCATCTGTCTATTTGCTTATCATCATAGTTG 2755
QY 2740 CTGACTCCCGGCGGTGATTAATCAAGATATACGAGATTAACGAAATTAACAGGCC 2799
Db 2756 CTGACTCCCGGCGGTGATTAATCAAGATATACGAGATTAACGAAATTAACAGGCC 2815
QY 2800 TGCATGTATACCGGAGAACCCACGCTCACCGGCTCCAGATTTATACGAAATTAACAGGCC 2859
Db 2816 TGCATGTATACCGGAGAACCCACGCTCACCGGCTCCAGATTTATACGAAATTAACAGGCC 2875
QY 2860 AGCCGAAAGGGCGAGGCGAGAAAGTGTCTGCAACTTATTCGCGCTCATCAGTCTAT 2919
Db 2876 AGCCGAAAGGGCGAGGCGAGAAAGTGTCTGCAACTTATTCGCGCTCATCAGTCTAT 2935
QY 2920 TAATTTGGCCGGGAAGCTAGATTAAGTGTGCGAGTTAATAGTTTCCGACAGTGT 2979
Db 2936 TAATTTGGCCGGGAAGCTAGATTAAGTGTGCGAGTTAATAGTTTCCGACAGTGT 2995
QY 2980 TGCATTGCTAACAGGCAATGCTGTGACGCTGCTGTTGTAATGCTTCAATCAGCTC 3039
Db 2996 TGCATTGCTAACAGGCAATGCTGTGACGCTGCTGTTGTAATGCTTCAATCAGCTC 3055
QY 3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAG 3099
Db 3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAG 3115
QY 3100 CTGCTTGGGCTCCGCAATGCTTGTCAAGATTAAGTTGGCCGAGTATCATCATAGT 3159
Db 3116 CTGCTTGGGCTCCGCAATGCTTGTCAAGATTAAGTTGGCCGAGTATCATCATAGT 3175
QY 3160 TATGAGACATGCAATATCTTACTGTATGCAATGCAATCCGTAAGTCTTTTCTGTGAC 3219
Db 3176 TATGAGACATGCAATATCTTACTGTATGCAATGCAATCCGTAAGTCTTTTCTGTGAC 3235
QY 3220 TGGTAGTACTTAACCAAGTATCTTGAATAGTATGCGGACGAGATTTCTGTTGAC 3279
Db 3236 TGGTAGTACTTAACCAAGTATCTTGAATAGTATGCGGACGAGATTTCTGTTGAC 3295
QY 3280 CCCGCGTCAATACCGGATTAATACCGGCGCATAGGAGAACTTTAAAGTCTCATCAT 3339
Db 3296 CCCGCGTCAATACCGGATTAATACCGGCGCATAGGAGAACTTTAAAGTCTCATCAT 3355

Oy 3340 TGAAGAACTCTTGGGGCGAAAACTCAAGATCTTACCGCTGTGATCCAGTTC 3399
Db 3356 TGAAGAACTCTTGGGGCGAAAACTCAAGATCTTACCGCTGTGATCCAGTTC 3415
Oy 3400 GATGTAACCACTCGTGCACCCAACTGATCTTCAAGCATCTTTACTTACACAGCGTTTC 3459
Db 3416 GATGTAACCACTCGTGCACCCAACTGATCTTCAAGCATCTTTACTTACACAGCGTTTC 3475
Oy 3460 TGGGTGACCAAAAAACAGAGAGCAAAATGCCGCAAAAAAGGAAATAGCGCGACAGGAA 3519
Db 3476 TGGGTGACCAAAAAACAGAGAGCAAAATGCCGCAAAAAAGGAAATAGCGCGACAGGAA 3535
Oy 3520 ATGTTGAATACATCACTCTCTTTTCAATATATTGAAGCATTTATCAGGGTTATG 3579
Db 3536 ATGTTGAATACATCACTCTCTTTTCAATATATTGAAGCATTTATCAGGGTTATG 3595
Oy 3580 TCTCATGAGCGGATATCATATTTGATGATTTAGAAAAATAAACAATAGGGGTTCCGCG 3639
Db 3596 TCTCATGAGCGGATATCATATTTGATGATTTAGAAAAATAAACAATAGGGGTTCCGCG 3655
Oy 3640 CACATTTTCCCGAAAAAGTCCAC 3662
Db 3656 CACATTTTCCCGAAAAAGTCCAC 3678

RESULT 12

AAQ13576
ID AAQ13576 standard; DNA; 5314 BP.

XX AAQ13576;

XX 02-DEC-1991 (first entry)

DE Plasmid pMTV1 containing 3 multiple cloning sites.

XX recombinant expression plasmid; phage phi X174; protein E;

KM bacterial ghost; ss.

OS Synthetic.

XX MO9113155-A.

XX 05-SEP-1991.

XX 24-FEB-1990; 90DE-04005874.

XX 24-FEB-1990; 90DE-04005874.

XX (BOEF) BOEHRINGER MANNHEIM GMBH.

XX Lubitz W, Szostak MP;

XX WPI; 1991-281471/38.

XX Carrier recombinant protein for vaccines against HIV, etc. - obtd. by

PT expressing fusion protein gene in gram negative bacteria and gene that

XX encodes for lytic membrane protein.

XX Example 10; Page 28-29; 45pp; German.

XX The carrier plasmid pMTV1 comprises the phage phi X174 E protein gene,

XX the ampicillin resistance gene, the lambda c1857 repressor gene and P(R)

XX from the lambda promoter/operator system. There are also lac operon

XX sequences and three multiple cloning sites. A foreign gene can be

XX inserted into the second multiple cloning site (mc82). The protein

XX encoded by the foreign gene is especially a viral antigenic peptide, e.g.

XX of HIV, HBV or EBV. See also AAQ13577 and AAQ13578

XX Sequence 5314 BP; 1438 A; 1276 C; 1310 G; 1290 T; 0 U; 0 Other;

Oy 20 ATTTTGTAAATTCGGTTAAATTTTGTAAATCAAGTCATTTTTTAACCAATAGGCC 79
Db 1651 ATTTTGTAAATTCGGTTAAATTTTGTAAATCAAGTCATTTTTTAACCAATAGGCC 1710
Oy 80 GAAATCGGCAAAATCCCTTATTAATCAAAAAGATAGACCGGATAGGGTTAGTGTG 139
Db 1711 GAAATCGGCAAAATCCCTTATTAATCAAAAAGATAGACCGGATAGGGTTAGTGTG 1770
Oy 140 CCAAGTTTGAAACAAAGATCCATTAAGAACGTGACCTTCAAGCTTCAAGGGCGAAA 199
Db 1771 CCAAGTTTGAAACAAAGATCCATTAAGAACGTGACCTTCAAGCTTCAAGGGCGAAA 1830
Oy 200 ACCGTTTATCAGGCGATGCGCCCTACGTATCAACATCACTTATCAAGTTTTTGGG 259
Db 1831 ACCGTTTATCAGGCGATGCGCCCTACGTATCAACATCACTTATCAAGTTTTTGGG 1890
Oy 260 TCGAGGTGCGTAAGCACTAAATCGGAACCTTAAGGGAGCCCGGTTTGAAGCTTGA 319
Db 1891 TCGAGGTGCGTAAGCACTAAATCGGAACCTTAAGGGAGCCCGGTTTGAAGCTTGA 1950
Oy 320 CGGGGAAAGCCGCGAAAGTGGCGAAGAAAGGAAAGGAAAGGAAAGGCGGCGCT 379
Db 1951 CGGGGAAAGCCGCGAAAGTGGCGAAGAAAGGAAAGGAAAGGAAAGGCGGCGCT 2010
Oy 380 AGGCGCTGCGCAATGTAAGCGGTACCGTGGCGTTAACACACACCCCGCGCTTAT 439
Db 2011 AGGCGCTGCGCAATGTAAGCGGTACCGTGGCGTTAACACACACCCCGCGCTTAT 2070
Oy 440 GCGCGCTGCAAGGCGGCGTCCATTTGCGCAATTCAGGCGGCACTGTGGGAAAGGGGA 499
Db 2071 GCGCGCTGCAAGGCGGCGTCCATTTGCGCAATTCAGGCGGCACTGTGGGAAAGGGGA 2130
Oy 500 TCGGTGCGGCGCTTGTGCTATTACGCGAGTGGCGAAAGGGGATGTGTCAAGCGGA 559
Db 2131 TCGGTGCGGCGCTTGTGCTATTACGCGAGTGGCGAAAGGGGATGTGTCAAGCGGA 2190
Oy 560 TTAAGTTGGGTAAACCGCAAGGTTTTCCAGTACGACGTTGTAAACGACCGCACTGAG 619
Db 2191 TTAAGTTGGGTAAACCGCAAGGTTTTCCAGTACGACGTTGTAAACGACCGCACTGAG 2250
Oy 620 CGCGCTGCTTATTCATTCAGTTTTTGAACCGGTGAGAGACCGGCGAGACTCGGCGCAAT 679
Db 2251 TTAAGTTGGGTAAACCGCAAGGTTTTCCAGTACGACGTTGTGTGAGT 2310
Oy 680 GTGTTTTACAGCGTATGAGAGCAAGATGATGCTGCACAGCTGCAGAACACCGAGCTA 739
Db 2311 ATGTTGCACTCTCAGTACAAATCTGCTGATGCCGCAATGTTAAGCCAGTATACACTC 2370
Oy 740 GATTAACTTGAAGAAATATCATATTTGTGACGTAGCTTAAAGATATCATGCGTAAAT 799
Db 2371 GATTAACTTGAAGAAATATCATATTTGTGACGTAGCTTAAAGATATCATGCGTAAAT 2430
Oy 800 TGAGCGATGGGATGTGTAATACGATCATATAGGGGAAATGGGTACCGGCGCCCT 859
Db 2431 TGAGCGATGGGATGTGTAATACGATCATATAGGGGAAATGGGTACCGGCGCCCT 2490
Oy 860 CGAGTCGACGATGCA-----TAAGCTTGAATGCAATTCCTGCAGCCCGGGGATC 913
Db 2491 CGAGTCGACGATGCAAGGTTTCAACGCTATACCGAAACGGCGAGAGAGTAAAGTTC 2550
Oy 914 CACTAAGTTTGAAGCGGCGGCAAGCGGTGAGCTTCAAGCTTGTGTTTCCCTTATGAGAG 973
Db 2551 CACTAAGTTTGAAGCGGCGGCAAGCGGTGAGCTTCAAGCTTGTGTTTCCCTTATGAGAG 2610
Oy 974 GGTAAATTAAGTCCAT-----GGCTCAATTTTGAAGAGATATTTCTGAGGTTAA 1026
Db 2611 GGTAAATTAAGTCCAT-----GGCTCAATTTTGAAGAGATATTTCTGAGGTTAA 2670
Oy 1027 TCTAGTCATCAGAGATCATATCGTGGGCTTTTTCGAGCTCAGTATCGCCAAAGCT 1086
Db 2671 TCTAGTCATCAGAGATCATATCGTGGGCTTTTTCGAGCTCAGTATCGCCAAAGCT 2730

Qy 1087 GGGCTATCTGGGCATCGGAGAGAAAGCCCGTCTTTCGCGAGGTTGAAGCGG 1146
| | | | |
Db 2731 CATGTTGAATGGCCCTTAAGAGATCCGGCAAGCTTGTGATGCTGCAAGTGCATTAAGA 2790
| | | | |
Qy 1147 CATGAAAGATTGTTCCGAGATGACTGTGCTGATGACGTTGAGGAAAGCAAGT 1206
| | | | |
Db 2791 GGAATCCCGACGCTCGAGCCGATTAATATGTTTCCGTAATTCAGCGCTTCATGAT 2850
| | | | |
Qy 1207 TTACCATGATGATGGGAAAGGTGTGGCATGACAGCCCTTAAACGTGAATCTGTGTG 1266
| | | | |
Db 2851 GAGACAGGCCGTTT--GAATGTGACGGGATGAACATTAATPAAGCAATGACGGCAATTA 2908
| | | | |
Qy 1267 AGGCCACTGGGATACAGTTCTGTGCGGCTTTTCGAGACAGCTTCGGATGCTAGCC 1326
| | | | |
Db 2909 AACTCAACAGG--AGCAGAAAGCAGAGGATATCCCAAAAGTCCAGCCGTACATAAACG 2965
| | | | |
Qy 1327 CGAAGCGCATGAGAACCCGAAACAATACCGGCGACAG-----CGGAAATGCGCGCGG 1380
| | | | |
Db 2966 CAAAGCTTAACGACGACGACGACGAGACGAGAGCGGTGAGTAGCAATCCAACTTTGTTACTC 3025
| | | | |
Qy 1381 GTGTGAGATTAAATGACAGCGGTGCGGCGCTGGGATTAACGTCAAGCAGAGACGGGAT- 1439
| | | | |
Db 3026 GTCAAGAAATCGAAATCATCTTCGGTTAAATCCAAAGCGCAGAAAGCTGAATGAGAAAT 3085
| | | | |
Qy 1440 ---CCTGGCTGAGATCCGCGCAAAATGGAATGGAATCCCGGTGAGTTACCGCGGGGGC 1496
| | | | |
Db 3086 CGAAGCTTAAGAGGGGGGGCCCGGTACCGAGCTTTTGTTCCTTTAGTAGAGGTTAATTCGA 3145
| | | | |
Qy 1497 GCTTGGGATTAATCATGCTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAATTC 1556
| | | | |
Db 3146 GCTTGGGATTAATCATGCTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAATTC 3205
| | | | |
Qy 1557 CACAACAATAGAGCGGAGACATTAAGTTAAAGCTGGGGTGCTTAATGAGTAGT 1616
| | | | |
Db 3206 CACAACAATAGAGCGGAGACATTAAGTTAAAGCTGGGGTGCTTAATGAGTAGT 3265
| | | | |
Qy 1617 AACTCATTAATTTGGTTGCGCTCATGCGCGCTTCAGTCCGGAACCTGTGCTGCC 1676
| | | | |
Db 3266 AACTCATTAATTTGGTTGCGCTCATGCGCGCTTCAGTCCGGAACCTGTGCTGCC 3325
| | | | |
Qy 1677 AACTCATTAATTTGGTTGCGCTCATGCGCGCTTCAGTCCGGAACCTGTGCTGCC 1736
| | | | |
Db 3326 AACTCATTAATTTGGTTGCGCTCATGCGCGCTTCAGTCCGGAACCTGTGCTGCC 3385
| | | | |
Qy 1737 CCGCTTCTCGCTCATGCTCATGCGCTCATGCGCTTCAGTCCGGAACCTGTGCTGCC 1796
| | | | |
Db 3386 CCGCTTCTCGCTCATGCTCATGCGCTCATGCGCTTCAGTCCGGAACCTGTGCTGCC 3445
| | | | |
Qy 1797 CTCACTCAAAAGCGGTTAATGCTTATCCACAGATCAGGGGATTAACGAGAAAGACA 1856
| | | | |
Db 3446 CTCACTCAAAAGCGGTTAATGCTTATCCACAGATCAGGGGATTAACGAGAAAGACA 3505
| | | | |
Qy 1857 TGTGAGCAAAAGCGCAGCAAAAGCGCAGAAACGTTAAAGCGCGGTTGCTGGCGTTT 1916
| | | | |
Db 3506 TGTGAGCAAAAGCGCAGCAAAAGCGCAGAAACGTTAAAGCGCGGTTGCTGGCGTTT 3565
| | | | |
Qy 1917 TCCATAGAGCTCCGCCCCCTGACAGACATCAAAATTCAGAGCTCAAGTCAAGAGTGGC 1976
| | | | |
Db 3566 TCCATAGAGCTCCGCCCCCTGACAGACATCAAAATTCAGAGCTCAAGTCAAGAGTGGC 3625
| | | | |
Qy 1977 GAAACCCGACAGGACTATAAGATACAGAGCGTTTCCCGTGGAGAGCTCCCTGCTGCT 2036
| | | | |
Db 3626 GAAACCCGACAGGACTATAAGATACAGAGCGTTTCCCGTGGAGAGCTCCCTGCTGCTG 2096
| | | | |
Qy 2037 CTCTGTTTCGAGCCCGGCTTACCGGATACCTGTCCGCTTCTCCCTCGGAAAGCG 2096
| | | | |
Db 3686 CTCTGTTTCGAGCCCGGCTTACCGGATACCTGTCCGCTTCTCCCTCGGAAAGCG 3745
| | | | |
Qy 2097 TGGCGTTTCTCATAGCTCAAGCTGTAGGTAATCAGATTGGTGGTGGTGGTGGTCCA 2156
| | | | |
Db 3746 TGGCGTTTCTCATAGCTCAAGCTGTAGGTAATCAGATTGGTGGTGGTGGTGGTCCA 3805
| | | | |
Qy 2157 AACTGGGCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAAC 2216
| | | | |

Db 3806 AGCTGGCTGTGTGACGAACCCCGTTACGCCCGACCGCTCGCTTATCCGGTAAC 3865
| | | | |
Qy 2217 ATGCTTGAATCCAAACCGGTAAAGACACGACTTAATGACCACTGGAGAGCAAGCTGTA 2276
| | | | |
Db 3866 ATGCTTGAATCCAAACCGGTAAAGACACGACTTAATGACCACTGGAGAGCAAGCTGTA 3925
| | | | |
Qy 2277 ACAGATTTAGACAGGAGGATGATGAGCGGCTCTCAAGATTTTGAAGTGGTGGCTTA 2336
| | | | |
Db 3926 ACAGATTTAGACAGGAGGATGATGAGCGGCTCTCAAGATTTTGAAGTGGTGGCTTA 3985
| | | | |
Qy 2337 ACTAGCGCTACATGAGAGACGATTTTGTATCTGCGCTGTGCGAAGCCAGTACT 2396
| | | | |
Db 3986 ACTAGCGCTACATGAGAGACGATTTTGTATCTGCGCTGTGCGAAGCCAGTACT 4045
| | | | |
Qy 2397 TCGGAAAAAGATTGGTAGCTCTGTAACCGGAAACAAACCAACCGCTGTAGCGGTGTT 2456
| | | | |
Db 4046 TCGGAAAAAGATTGGTAGCTCTGTAACCGGAAACAAACCAACCGCTGTAGCGGTGTT 4105
| | | | |
Qy 2457 TTTTGTGTTGACAGACGAGATTACCGCGACAGAAAAAGATCTCAAGAGATCTTTGA 2516
| | | | |
Db 4106 TTTTGTGTTGACAGACGAGATTACCGCGACAGAAAAAGATCTCAAGAGATCTTTGA 4165
| | | | |
Qy 2517 TCTTTTCTACGGGCTGACAGCTGAGTGGAAAGAAAACTACGTTAAGGATTTGGTCA 2576
| | | | |
Db 4166 TCTTTTCTACGGGCTGACAGCTGAGTGGAAAGAAAACTACGTTAAGGATTTGGTCA 4225
| | | | |
Qy 2577 TGAATTAATCAAAAGATCTTCACTAGATCTTTTAAATTAAGATTTTAAAT 2636
| | | | |
Db 4226 TGAATTAATCAAAAGATCTTCACTAGATCTTTTAAATTAAGATTTTAAAT 4285
| | | | |
Qy 2637 CAATCAAGATTAATTAAGTAAGTGTGACAGTTGCAATGCTTAATCAAGTAGAG 2696
| | | | |
Db 4286 CAATCAAGATTAATTAAGTAAGTGTGACAGTTGCAATGCTTAATCAAGTAGAG 4345
| | | | |
Qy 2697 CACCAATTCAGGCACTTGTCTAATTTGTTATCTCAATGTTGCTGACCTCCGCTGCT 2756
| | | | |
Db 4346 CACCAATTCAGGCACTTGTCTAATTTGTTATCTCAATGTTGCTGACCTCCGCTGCT 4405
| | | | |
Qy 2757 AGATTAATCAAGATTCAGGAGGCTTACCAATGAGCCCGCAGTGTGATGATACCGGAG 2816
| | | | |
Db 4406 AGATTAATCAAGATTCAGGAGGCTTACCAATGAGCCCGCAGTGTGATGATACCGGAG 4465
| | | | |
Qy 2817 ACCAGCGCTACCGGCTCAAGATTTATCAGCAATTAACCAAGCAGCGGAGGCGGAGC 2876
| | | | |
Db 4466 ACCAGCGCTACCGGCTCAAGATTTATCAGCAATTAACCAAGCAGCGGAGGCGGAGC 4525
| | | | |
Qy 2877 GCAAGATGCTCTGCACTTATCCGCTCCATCAAGTCTAATTAATTTGTCGGGAG 2936
| | | | |
Db 4526 GCAAGATGCTCTGCACTTATCCGCTCCATCAAGTCTAATTAATTTGTCGGGAG 4585
| | | | |
Qy 2937 CTGAGATTAAGTGTGCTGCAATTAATGTTGCGCAAGTGTGCGCATGCTCAAGGCA 2996
| | | | |
Db 4586 CTGAGATTAAGTGTGCTGCAATTAATGTTGCGCAAGTGTGCGCATGCTCAAGGCA 4645
| | | | |
Qy 2997 TGTGATGCTACGCTGCTGTTGGTGAAGCTTCAATCAAGTCCGTTCCAAAGATCA 3056
| | | | |
Db 4646 TGTGATGCTACGCTGCTGTTGGTGAAGCTTCAATCAAGTCCGTTCCAAAGATCA 4705
| | | | |
Qy 3057 GGCAGTTTACATGATCCCGCATGTTGTGCAAAAGAGGTTAGTCTCTGCGCTCCCA 3116
| | | | |
Db 4706 GGCAGTTTACATGATCCCGCATGTTGTGCAAAAGAGGTTAGTCTCTGCGCTCCCA 4765
| | | | |
Qy 3117 TCGTGTGAGAAAGTGGGCGCAAGTGTATCACTCATGTTATGAGCAGACTGACATA 3176
| | | | |
Db 4766 TCGTGTGAGAAAGTGGGCGCAAGTGTATCACTCATGTTATGAGCAGACTGACATA 4825
| | | | |
Qy 3177 ATTCTCTTAATGTCATGCAATCCGTAAAGTCTTTCTGATGAGTGGTGAATCTCAACCA 3236
| | | | |
Db 4826 ATTCTCTTAATGTCATGCAATCCGTAAAGTCTTTCTGATGAGTGGTGAATCTCAACCA 4885
| | | | |
Qy 3237 AGTCATCTGAGAAATGATGATCCGGGACCAAGTTGCTTTGCCGGGCTCAATACGGG 3296
| | | | |

QY 837 GAATTTGGATTACCGGGGCCCCCTCGAGGTGACGGTATTCGATTAATTCGAAATTC 896
| | | | |
Db 840 GGACAGGGCCATCGCAATTCGAAATTTGTTGTAATAGTCTGCTAGTTGAACGCTTC 899
| | | | |
QY 897 CTGACGCCGGGGGGAATTCACTAGTT----CTAGACGGGCCGCAACCGGGGTGAGCTCCA 952
| | | | |
Db 900 CATTCTTAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCATTTCTTTGTTGTG 959
| | | | |
QY 953 GCTTTTGTCCCTTTAGTGAAGGTTAATTAGATCCATGCGTCAATTTTAAAGCAGACTAT 1012
| | | | |
Db 960 CCATATGATATACATTTGTGTAGTTATAGTTGATTCCAATTTGTGTCAGAAATGTTTC 1019
| | | | |
QY 1013 CTTTCTAGGGTTAATCTAGCTGCATCAGAGATCATTCGTGGGCTTTTTCGGCTGAG 1072
| | | | |
Db 1020 CATCTCTTTTAAATCAATACCTTTTAACTGATTCATTAACAAGGATACCTTCA 1079
| | | | |
QY 1073 TCATGCCCAAGCTGCGCTATCTGCGCATCGGGAGAGAAAGCCGTCCTTTTCCCG 1132
| | | | |
Db 1080 ACTTGACTTCAAGCAGTGTCTTGTAGTTCCCGTATCTTTGAAATAATATGTTCTTTCCT 1139
| | | | |
QY 1133 CGAGGTGAAGCGGATGAGAAAGATTGCGAGGATGACTGTGCTGCATTGACGTTGA 1192
| | | | |
Db 1140 GTACATACTTCGGGCATGCGACTCTTGAATAAGTCAATGCCGTTCAATGATCCGGGT 1199
| | | | |
QY 1193 GCGAAACGCACTTTTACATGATGATTCGGAAAGTGTGGCCATGACACGCTTTAACG 1252
| | | | |
Db 1200 ATCTTGAATAACATTGAACCACTAAGAAAGTATGACAAAGTGTGGCCATGAAACAG 1259
| | | | |
QY 1253 TGAATCTGCTTCAAGGCCACCTGGGAATACCAATTCGTGCGGCTTTTCGGAACAGTT 1312
| | | | |
Db 1260 GTAGTTTTCAGTATGTGCAATTAATTTAAGGTAATTTCCGATGTTGCACTTCCT 1319
| | | | |
QY 1313 CCGATGAGTCAAGCCGGAAGCGCATCAGAACCCGAAACATACCGGCGACAGCCGAACTG 1372
| | | | |
Db 1320 CACCTCTCCACATCAAGAAATTTGTGCCATTTAACTAATCACTAATTAATCAACAGAA 1379
| | | | |
QY 1373 CCGTGC-----GGTGTCAATTAATGACAGCGGTGCGGCTGCGGATAT 1418
| | | | |
Db 1380 TTGGGAACAATCCAGTAAAGATTCTTCCTTTGCTAGCCATTTCTTGGCCATGAAAT 1439
| | | | |
QY 1419 TACGTCAAGCAGAGCGGGATTCCTGCGTGAATGCCGAGAAATGACATGATACCCCGT 1478
| | | | |
Db 1440 TCTGTCAGCCCGGGGGATTCACATGATTCAGAGCGCGCCACCGGCGAGACTCCACAG 1499
| | | | |
QY 1479 GAGTTACCGGCGGG-----CGCGCTTGGCGTAACTAGTGTCAATGCGTTT 1525
| | | | |
Db 1500 TTTTGTTCCTTTAGTAGAGGTTAATTCGAGCTTGGCGTATCATGTCTATGCTGTTT 1559
| | | | |
QY 1526 CCTGTGTGAATTTGTTATCCGCTCACAATTCACAACAATACGAGCCGGAAACATTAAG 1585
| | | | |
Db 1560 CCTGTGTGAATTTGTTATCCGCTCACAATTCACAACAATACGAGCCGGAAACATTAAG 1619
| | | | |
QY 1586 TGTAAAGCTGGGGTGCCTAATGATGAGTAACTCAATTAATTCGTTGGCTCACTG 1645
| | | | |
Db 1620 TGTAAAGCTGGGGTGCCTAATGATGAGTAACTCAATTAATTCGTTGGCTCACTG 1679
| | | | |
QY 1646 CCGGCTTTTCCAGTTCGGGAAACCTGTGCTGCAGCTCATTAATGATTCGGCCAAACGCGCG 1705
| | | | |
Db 1680 CCGGCTTTTCCAGTTCGGGAAACCTGTGCTGCAGCTCATTAATGATTCGGCCAAACGCGCG 1739
| | | | |
QY 1706 GGGAGAGGGGTTTGGGTAATGGGCGCTCTTCGCGTTCTCGCTCACTGATTCGCTGCGC 1765
| | | | |
Db 1740 GGGAGAGGGGTTTGGGTAATGGGCGCTCTTCGCGTTCTCGCTCACTGATTCGCTGCGC 1799
| | | | |
QY 1766 TCGGTGTTTCGCTGCGGAGCGGATACGCTCACTCAAGGCGGTAATAGGTTATTC 1825
| | | | |
Db 1800 TCGGTGTTTCGCTGCGGAGCGGATACGCTCACTCAAGGCGGTAATAGGTTATTC 1859
| | | | |
QY 1826 ACGAATTCAGGGGATTAACGAGAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1885
| | | | |
Db 1860 ACGAATTCAGGGGATTAACGAGAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1919
| | | | |

QY 1886 AACGTAATAAGGCGCGGTTGCTGAGGTTTTTTCATAGTCTCGCCCCCTGACGAGCAT 1945
| | | | |
Db 1920 AACGTAATAAGGCGCGGTTGCTGAGGTTTTTTCATAGTCTCGCCCCCTGACGAGCAT 1979
| | | | |
QY 1946 CACAAAATTCAGGCTCAAGTCAAGAGTGGCGAAACCCGACAGACTATTAAGATACAG 2005
| | | | |
Db 1980 CACAAAATTCAGGCTCAAGTCAAGAGTGGCGAAACCCGACAGGACTATTAAGATACAG 2039
| | | | |
QY 2006 GCGTTTCCCGCTGGAAGCTCCCTCGGCGCTCTCTGTTTCCGACCTGCGCTTACCGGA 2065
| | | | |
Db 2040 GCGTTTCCCGCTGGAAGCTCCCTCGGCGCTCTCTGTTTCCGACCTGCGCTTACCGGA 2099
| | | | |
QY 2066 TACCTGTCCGCTTTCTCCCTTCGGGAACGTGCGCTTCTCATAGTCAAGCTGTAG 2125
| | | | |
Db 2100 TACCTGTCCGCTTTCTCCCTTCGGGAACGTGCGCTTCTCATAGTCAAGCTGTAG 2159
| | | | |
QY 2126 TATCTCAATTCGGTGTAGTCCGTTGCTCAAGCTGAGGCTGTGTGACAGAACCCCGTT 2185
| | | | |
Db 2160 TATCTCAATTCGGTGTAGTCCGTTGCTCAAGCTGAGGCTGTGTGACAGAACCCCGTT 2219
| | | | |
QY 2186 CAGCCCGAACCGCTGCGCTTATCCGGTAATATGCTTTGAGTCCAAACCCGGTAAGAC 2245
| | | | |
Db 2220 CAGCCCGAACCGCTGCGCTTATCCGGTAATATGCTTTGAGTCCAAACCCGGTAAGAC 2279
| | | | |
QY 2246 GACTTATCCCACTGGCAGCAGCCACTGTGTAAACAGATTAAGCAGACGAGTATAGGC 2305
| | | | |
Db 2280 GACTTATCCCACTGGCAGCAGCCACTGTGTAAACAGATTAAGCAGACGAGTATAGGC 2339
| | | | |
QY 2306 GGTGTACAGAGTTCCTGAAGTGTGGGCTTAATCCGCTACACTAAGAAAGCAATATT 2365
| | | | |
Db 2340 GGTGTACAGAGTTCCTGAAGTGTGGGCTTAATCCGCTACACTAAGAAAGCAATATT 2399
| | | | |
QY 2366 GGTATCTGCGCTCTGCTGAAGCAAGTACTCTTCGAAATAAGTTGATGCTTTGATCC 2425
| | | | |
Db 2400 GGTATCTGCGCTCTGCTGAAGCAAGTACTCTTCGAAATAAGTTGATGCTTTGATCC 2459
| | | | |
QY 2426 GGCAAAACAAACCAACGCTGTGTAGCGGTGTTTTTTTGTTCGAAAGCAGAGATTAACGCGC 2485
| | | | |
Db 2460 GGCAAAACAAACCAACGCTGTGTAGCGGTGTTTTTTTGTTCGAAAGCAGAGATTAACGCGC 2519
| | | | |
QY 2486 AGAAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCACTGG 2545
| | | | |
Db 2520 AGAAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCACTGG 2579
| | | | |
QY 2546 AAGCAAACTACAGTTAAGGATTTTGTCTATGAGATTAATCAAAAGATCTTCACTAG 2605
| | | | |
Db 2580 AAGCAAACTACAGTTAAGGATTTTGTCTATGAGATTAATCAAAAGATCTTCACTAG 2639
| | | | |
QY 2606 ATCTTTTAATTTAAAAATGAAGTTTAAATCAATCTTAAAGTATATAGATAACTTGG 2665
| | | | |
Db 2640 ATCTTTTAATTTAAAAATGAAGTTTAAATCAATCTTAAAGTATATAGATAACTTGG 2699
| | | | |
QY 2666 TCTGACAGTTAACATGCTTAATCAGTAGGCACTTATCAAGGATCTGTCTAATTCGT 2725
| | | | |
Db 2700 TCTGACAGTTAACATGCTTAATCAGTAGGCACTTATCAAGGATCTGTCTAATTCGT 2759
| | | | |
QY 2726 TCAATCAATGTTGCTGACTCCCGTGTGTAGTAACTAGATCGGAGGGCTTAACA 2785
| | | | |
Db 2760 TCAATCAATGTTGCTGACTCCCGTGTGTAGTAACTAGATCGGAGGGCTTAACA 2819
| | | | |
QY 2786 TCTGAGCCCAAGTGTGAGATGATACGCGAGAACCAACGCTCAACCGGCTCAGATTATCA 2845
| | | | |
Db 2820 TCTGAGCCCAAGTGTGAGATGATACGCGAGAACCAACGCTCAACCGGCTCAGATTATCA 2879
| | | | |
QY 2846 GCAATTAACCAAGCCAGCGGAGGCGGAGCGAGAAAGTGTCTGCAACTTTATCCGCC 2905
| | | | |
Db 2880 GCAATTAACCAAGCCAGCGGAGGCGGAGCGAGAAAGTGTCTGCAACTTTATCCGCC 2939
| | | | |
QY 2906 TCCATCAGTCTATTAATGTTGCTCCGGGAAGCTAAGAGTATGTTCCGCAATTAATAGT 2965
| | | | |
Db 2940 TCCATCAGTCTATTAATGTTGCTCCGGGAAGCTAAGAGTATGTTCCGCAATTAATAGT 2999
| | | | |
QY 2966 TTGCGCAACGTTGTGCAATTCCTACAGGCATGTGTGTCAACGCTGTGTTGTGTATG 3025
| | | | |

```

Db      3000 TTGGGCAAGCTGTGGCCATTGCTACAGGCACTCGTGTCTACGCTGCTGTTGGTATG 3059
Qy      3026 GCTTCATTACAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCAATGTTGTGC 3085
Db      3060 GCTTCATTACAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCAATGTTGTGC 3119
Qy      3086 AAAAAGGGGTTAGCTCTCGTCCGTCCGATCCGTTGTCAGAAAGTAAGTTGGCCGCACTG 3145
Db      3120 AAAAAGGGGTTAGCTCTCGTCCGTCCGATCCGTTGTCAGAAAGTAAGTTGGCCGCACTG 3179
Qy      3146 TTATCACTCATAGTGTATGAGCACTGATATATCTCTTAATGTCATGCAATCCGTAGA 3205
Db      3180 TTATCACTCATAGTGTATGAGCACTGATATATCTCTTAATGTCATGCAATCCGTAGA 3239
Qy      3206 TGCCTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTGAGAAATGTAATGCGCGCA 3265
Db      3240 TGCCTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTGAGAAATGTAATGCGCGCA 3299
Qy      3266 CCGAGTGTCTGTGCGCGCGCTCAATAGGGAATAACCGCGCCCATAGCAAGAACTTTA 3325
Db      3300 CCGAGTGTCTGTGCGCGCGCTCAATAGGGAATAACCGCGCCCATAGCAAGAACTTTA 3359
Qy      3326 AAAGTGTCACTCATTTGAAACAGTTCTTCCGCGCGCAAACTCTCAAGGATCTTACCGCTG 3385
Db      3360 AAAGTGTCACTCATTTGAAACAGTTCTTCCGCGCGCAAACTCTCAAGGATCTTACCGCTG 3419
Qy      3386 TTGAGATCAAGTTCGATGTAAACCACTGTGTGCAACCACTGATCTTTCAGCATCTTTTACT 3445
Db      3420 TTGAGATCAAGTTCGATGTAAACCACTGTGTGCAACCACTGATCTTTCAGCATCTTTTACT 3479
Qy      3446 TTCCACCAAGCTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATGCCGCAAAAAGGGAATA 3505
Db      3480 TTCCACCAAGCTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATGCCGCAAAAAGGGAATA 3539
Qy      3506 AGGCGACACAGGAAATGTTGAATCTCATACCTCTCTTTTCAATATTTATTTGAAGCAAT 3565
Db      3540 AGGCGACACAGGAAATGTTGAATCTCATACCTCTCTTTTCAATATTTATTTGAAGCAAT 3599
Qy      3566 TATCAGAGGTTATTTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATTAACA 3625
Db      3600 TATCAGAGGTTATTTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATTAACA 3659
Qy      3626 ATAGGGGTTCCGCGCACTTTCCCGAAGAAAGTGCAC 3662
Db      3660 ATAGGGGTTCCGCGCACTTTCCCGAAGAAAGTGCAC 3696

```

RESULT 14

```

AAD27066
ID      AAD27066 standard; DNA; 4001 BP.
XX
AC      AAD27066;
XX
DT      09-APR-2002 (first entry)
XX
DE      Plasmid pGN59A DNA.
XX
KM      Vector construct; RNA inhibition; RNAi; gene expression control;
XX      pGN49A plasmid; ds.
XX
OS      Unidentified.
XX
XX      WO20018121-A1.
XX      22-NOV-2001.
XX      18-MAY-2001; 2001WO-IB001068.
XX      19-MAY-2000; 2000GB-00012233.
XX      (DEVG-) DEVGEN NV.
XX

```

```

PI      Plactinck G, Renard J, Bogaert T;
XX
DR      WPI; 2002-121984/16.
XX
PT      A new DNA vector construct containing opposable promoter and terminator
PT      sequences flanking a cloning site are useful for the expression of double
PT      stranded RNA useful for inhibition of RNA in gene expression control.
XX
PS      Claim 24; Fig 15; 75pp; English.
XX
CC      The present invention relates to improved vector constructs comprising
CC      two promoters in opposite orientation to each other, an inter-promoter
CC      region downstream of the 3' end of both promoters, a cloning site in the
CC      inter-promoter region and a transcription terminator downstream of the 3',
CC      end of the first promoter and the cloning site and operably linked to the
CC      first promoter. The constructs of the invention and the bacteria
CC      harboring the constructs are used to produce double stranded RNA for RNA
CC      inhibition (RNAi) and can be used as a tool for controlling gene
CC      expression. The present sequence is pGN59A plasmid DNA
XX
SQ      Sequence 4001 BP; 1034 A; 967 C; 1030 G; 970 T; 0 U; 0 Other;
XX
Query Match      71.3%; Score 2609.4; DB 6; Length 4001;
Best Local Similarity 83.1%; Pred. No. 0;
Matches 3069; Conservative 0; Mismatches 591; Indels 35; Gaps 7;
Qy      1 CTAATTTGTAAAGCGTTAATATTTTGTAAATTTGCGGTTAAATTTTGTAAATCAAGCTC 60
Db      63 CGAAATGTAAACGTTAATATTTTGTAAATTTGCGGTTAAATTTTGTAAATCAAGCTC 122
Qy      61 ATTTTAAACCAATAGCCGCAAAATCCCTTTAATCAAAAGAAATAGACCGA 120
Db      123 ATTTTAAACCAATAGCCGCAAAATCCCTTTAATCAAAAGAAATAGACCGA 182
Qy      121 GATAGGGTGAAGTGTGTTCCAGTTTGAACCAAGTCCACTATTTAAAGAAAGTGAAGTC 180
Db      183 GATAGGGTGAAGTGTGTTCCAGTTTGAACCAAGTCCACTATTTAAAGAAAGTGAAGTC 242
Qy      181 CAACGTCAAAAGGGGCAAAAACGTCATACAGGGGAGTGGCCCACTAGCGTAACATCAAC 240
Db      243 CAACGTCAAAAGGGGCAAAAACGTCATACAGGGGAGTGGCCCACTAGCGTAACATCAAC 302
Qy      241 CTAATCAAGTTTGTGGGTGAGAGTCCGTTAAAGCACTAAATGGAACCTTAAGGAG 300
Db      303 CAAATCAAGTTTGTGGGTGAGAGTCCGTTAAAGCTCTAAATGGAACCTTAAGGAG 362
Qy      301 CCCCCGATTTAGACTTGAACGGGGAAGACGGCGCAACGTGGCGAAGAAAGGAAGAA 360
Db      363 CCCCCGATTTAGACTTGAACGGGGAAGACGGCGCAACGTGGCGAAGAAAGGAAGAA 422
Qy      361 AGCGAAAGAGCGGGGCTAGAGGGCTGCGCAAGTGTAGCGGTCAACGCTGCGCTAACAC 420
Db      423 AGCGAAAGAGCGGGGCTAGAGGGCTGCGCAAGTGTAGCGGTCAACGCTGCGCTAACAC 482
Qy      421 CACACCCGCGCGCTTAATGCGCGCTACAGGGGCGCTCCATTCGCTACGCTGCGG 480
Db      483 CACACCCGCGCGCTTAATGCGCGCTACAGGGGCGCT -CCATTCGCGCTACGCTGCGG 541
Qy      481 CAACTGTTGGGAAGGGGAGTCGAGGCGCTCTTCTGCTATTACGCAAGCTGGGAAAG 540
Db      542 CAACTGTTGGGAAGGGGAGTCGAGGCGCTCTTCTGCTATTACGCAAGCTGGGAAAG 601
Qy      541 GGGATGTGTCGAAGGCGATTAAGTTGGGTAAACGCCAGGTTTTTCCAGTCAAGCGTTG 600
Db      602 GGGATGTGTCGAAGGCGATTAAGTTGGGTAAACGCCAGGTTTTTCCAGTCAAGCGTTG 661
Qy      601 TAAACGACGCGCCAGTAAAGCGCGCT-----CGTTCACTTCAAGTTTGTGAA 646
Db      662 TAAACGACGCGCCAGTAAATTTGAATACGATCACTAATAGGCGCAATTTGAGCTCGGTAC 721
Qy      647 CCGGTGAGAGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGCGGTAGAGCAGATG 706
Db      722 CCGGGGATCCTCTTAGAGATCCTTCGACCTGAGATCATTTGTGCTGAAAGAGATCTGGAT 781

```


QY 1707 AAGATGCTCGACGCTGCGAGAACGACGAGTAACTAGAT-----TAACTTGAAGAAATGA 759
Db 782 CCGGCTTACTAATAAGCCAGATTAACATATGCTATTTGGCGGCTGATTTTGGGGATTA 841
QY 760 TCATATATGTTGACAGTACGTTTAAAGATTAATCAATGCTTAATTTG-----ACGATGGGAT 812
Db 842 GAATATATATGATATATATATCCGAAAGTATGTCAAAAAGAGGTGTGCTATGAAAGCAGCG 901
QY 813 CTGTAATACGACTCACTATATAGGGGAAATTGGGTACCGGGCCCCCTCGAGGTCAAGCGT 872
Db 902 TATTAACGTACAGATTGACAGCGACAGCTATCAGTTGTCTAAGGCATATATGATGCAAT 961
QY 873 ATCGATAGCTTGAATATCGAATTCGTAAGCCCGGGGAGTCCACTATGTTCTAGAGCGGC 932
Db 962 ATCTCGGCTGTGTAAGCACACATGCAAGAAATGAAGCCCGTGTCTGTCCGACCGACGC 1021
QY 933 GCCACCGCGGTGAGCTCCAGCTTTGTTCCCTTTAGTAGGGTTAATTAATCCCATGC 992
Db 1022 TGAAGAGCGGAAATACAGAAAGGAGTGGCTGAGGTGCCCGGTTTATGAATGAACGAC 1081
QY 993 GTCAATTTTACGAGACTA---TCTTTAGGGTTATCTAGCTGCATCAGATCATATC 1049
Db 1082 TCTTTTGCTGACGAGAACAGGACTGTGTAAATGCAGTTTAAAGTTTACACTATTAAG 1141
QY 1050 GTCCGGCTCTTTTTCGGGCTCACTCATCCGCAAGCTGAGCTATCTGGGCAATCGGGAG 1109
Db 1142 AGAGAGCCGTATTCGTCTGTTGTGGATGTACAGATGATATATTGACAGCCCGGGCG 1201
QY 1110 GAAGAGCCCGTGGCTTTTCCCGCAGAGTTGAAGCGCATGGAAGATTTGCCAGAGAT 1169
Db 1202 ACGGATGTGATCCCTCGGACC-AGTGAACGTCTCTTAAAGGATTAAGTCCCGTGAAC 1260
QY 1170 GACTGCTGCTG--CATTAAGTTGAGCGAAAAACGACGTTTACATGATGATTTGGGAA 1227
Db 1261 TTTACCCGATGTGCAATTCGGGAGTAAAGCTGGCGCATGATGACACCGATATGGCCA 1320
QY 1228 GTGTGGCATGACGCTTTTAAAGCTGAATGTTTGTTCAGGCCACCTGGGATACAGTT 1287
Db 1321 GTGTCCGCTCTCCCTTATCGGGAGAAAGTGGCTGATCTCAGCCACCGGAAATGACA 1380
QY 1288 CGTCCGGGCTTTTCCGACACAGTTCGCGATGTCACGCCGAAAGCGATACGACACCGA 1347
Db 1381 TCAAAAACGCCATTAACCTGATGTTCTGGGAAATTAATATGTCAGGCTCCCTTAATACA 1440
QY 1348 ACAATACCGGCGAGACGCGGAACCTGCGGTGTCAGATTATGACAGCGGTGGG 1407
Db 1441 GCTTTTCAGGACAAATGATCTCGAGGATCTTCATACCTACCAAGTTCTGGCCCTGACG 1500
QY 1408 CGTGGGATATTAAGTACGAGGAGACGGGTATCTCGGTGAGATGCGCGAAGATGCAAT 1467
Db 1501 GTCCCGGCGGAGCTCTAGAGTCCGAAAGCTTCCGCCCTTATAGTGAAGTCGTATTAAG 1560
QY 1468 GGAATACCCGGAATTACCGGCGGGCGGCTTGGGCTATCATGTCATAGCTGTTTTC 1527
Db 1561 CTGAGATTTCTATAGTCACTTAATAGCTTGGGCTATATCATGCTAGCTGTTTTC 1620
QY 1528 TGTGTGAATTTGTATCCGCTCACAATTCACACAACATACAGACCGGAGCATTAAGT 1587
Db 1621 TGTGTGAATTTGTATCCGCTCACAATTCACACAACATACAGACCGGAGCATTAAGT 1680
QY 1588 TAAAGCTGGGGTCCCTAATAGTAGTACTAATCAATTAATTTGGTTGCGTCACTGCC 1647
Db 1681 TAAAGCTGGGGTCCCTAATAGTAGTACTAATCAATTAATTTGGTTGCGTCACTGCC 1740
QY 1648 CGCTTTCCAGTCCGGAACCTGTGCGGACAGTGCATTAATATATGCGCAACGCGCGG 1707
Db 1741 CGCTTTCCAGTCCGGAACCTGTGCGGACAGTGCATTAATATATGCGCAACGCGCGG 1800
QY 1708 GAGAGCGGTTTGTGATTTGGCGCTCTTCCGCTTCTCGCTCATGATGATGCGCTC 1767
Db 1801 GAGAGCGGTTTGTGATTTGGCGCTCTTCCGCTTCTCGCTCATGATGATGCGCTC 1860

QY 1768 GGTGCTTCGCTGCGGAGAGCGGTATCAGTCACTCAAGAGCGGTAATACGTTATCCAC 1827
Db 1861 GGTGCTTCGCTGCGGAGAGCGGTATCAGTCACTCAAGAGCGGTAATACGTTATCCAC 1920
QY 1828 AGAATCAGGGGATTAACCAAGAAAGAAATGTGACAAAGAGCCAGCAAAAGGCGAGAA 1887
Db 1921 AGAATCAGGGGATTAACCAAGAAAGAAATGTGACAAAGAGCCAGCAAAAGGCGAGAA 1980
QY 1888 CCGTAAAGAGCGCGGTGTGCGGCTTTTTCATAGGCTCCGCCCCCTGACGAGATCA 1947
Db 1981 CCGTAAAGAGCGCGGTGTGCGGCTTTTTCATAGGCTCCGCCCCCTGACGAGATCA 2040
QY 1948 CAAAAATCGACGCTCAAGTATAGAGTGGGAAACCCGACAGACTAATAAGATACAGGC 2007
Db 2041 CAAAAATCGACGCTCAAGTATAGAGTGGGAAACCCGACAGACTAATAAGATACAGGC 2100
QY 2008 GTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGGCGCTTAACGGATA 2067
Db 2101 GTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGGCGCTTAACGGATA 2160
QY 2068 CCGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGATAGTA 2127
Db 2161 CCGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGATAGTA 2220
QY 2128 TCTCAATTGAGTATAGTGTGCTCCGCAAGCTGGGCTGTGACGAAACCCCGTTCA 2187
Db 2221 TCTCAATTGAGTATAGTGTGCTCCGCAAGCTGGGCTGTGACGAAACCCCGTTCA 2280
QY 2188 GCCCGACCGCTGCGCTTATCCGTAACCTATGTTAGTTCGCAACCCGATPACAGCA 2247
Db 2281 GCCCGACCGCTGCGCTTATCCGTAACCTATGTTAGTTCGCAACCCGATPACAGCA 2340
QY 2248 CTTATCGCACTGAGCAGAGCACTGTAACAGATTAAGCAGACGAGTATGTAAGCGG 2307
Db 2341 CTTATCGCACTGAGCAGAGCACTGTAACAGATTAAGCAGACGAGTATGTAAGCGG 2400
QY 2308 TGTCTCAAGTCTTGAAGTGTGGCTTAACGAGCTACACTGAGAGCAAGTATTTGG 2367
Db 2401 TGTCTCAAGTCTTGAAGTGTGGCTTAACGAGCTACACTGAGAGCAAGTATTTGG 2460
QY 2368 TATCTGCGCTGCTGGAAGCGATTAACCTCGAAAAAGAGTTGTAAGCTTGTGATCCG 2427
Db 2461 TATCTGCGCTGCTGGAAGCGATTAACCTCGAAAAAGAGTTGTAAGCTTGTGATCCG 2520
QY 2428 CAACCAACACCGCTGTGACGGTGTGTTTGTGTTGCAACAGAGATTACGGGAG 2487
Db 2521 CAACCAACACCGCTGTGACGGTGTGTTTGTGTTGCAACAGAGATTACGGGAG 2580
QY 2488 AAAAAAGATCTCAAGAAAGATCTTGTGATCTTTTCTAGGGGTCTGACGCTGATGAA 2547
Db 2581 AAAAAAGATCTCAAGAAAGATCTTGTGATCTTTTCTAGGGGTCTGACGCTGATGAA 2640
QY 2548 CGAAACTCAGTTAAGGATTTTGTGATAGATTATCAAAAAGGATCTTCACTAGAT 2607
Db 2641 CGAAACTCAGTTAAGGATTTTGTGATAGATTATCAAAAAGGATCTTCACTAGAT 2700
QY 2608 CTTTTTAATTAATAATGAAGTTTAAATCAATCTAAGATATATAGATAACTTGTGTC 2667
Db 2701 CTTTTTAATTAATAATGAAGTTTAAATCAATCTAAGATATATAGATAACTTGTGTC 2760
QY 2668 TGAACGTTACCAATGCTTATCAGTAGGACCTTATTCAGAGATCTGTCTATTTTCGTT 2727
Db 2761 TGAACGTTACCAATGCTTATCAGTAGGACCTTATTCAGAGATCTGTCTATTTTCGTT 2820
QY 2728 ATCCATAGTTGCTGACTCCCGTGTGATTAACCTACATACGAGAGGCTTAACATC 2787
Db 2821 ATCCATAGTTGCTGACTCCCGTGTGATTAACCTACATACGAGAGGCTTAACATC 2880
QY 2788 TGGCCCAAGTGTGCAATATACGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC 2847
Db 2881 TGGCCCAAGTGTGCAATATACGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC 2940
QY 2848 AATTAACAGCCAGCGGAGCGGAGCGGAGAGTGTCTCTCAACTTTATCCGCTC 2907

QY 601 TAAACGACGCGCAGTGAAGCGCGCTGTTCAATTCACGTTTGTGAACCCGTGAGAGACGG 660
Db 662 TAAACGACGCGCAGT----- 677
QY 661 GCAAGCTCGCGGTCAAAATGTGTTTACAGCGTGATGAGACAGATGAAGATGCTCGACAC 720
Db 678 ----- 677
QY 721 GCTGACAGACGCGAGCTAGATTACCTTAGAAAAATATCATATTTGTGACGTAA 780
Db 678 ----- 677
QY 781 AGATTAATCATGCGTAAATTAATGACGATGGGATCTGTAAATACACTCATATAGGGCAAT 840
Db 678 -----GAATTTGAATACACTCATATAGGGCAAT 708
QY 841 TTGGGTAACGGGCCCCCTCGAGGTGACCGGTATCGATAAGCTTGATATCGAATTC--CT 898
Db 709 TCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCAAGGGGTTATGCTAGTGAATTTCTGCA 768
QY 899 GCACCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGAGCTCCAGC---T 955
Db 769 GGGTACCCGGGATCTCTAGAGATCCCTGACCTGAGATCATTTGTGCGCGCGAT 828
QY 956 TTTGTTCCCTTTAGTAGGGTTAATTAATGATCCCATGCGCAATTTTACGACAGATATTT 1015
Db 829 TCTTATCACTGATAGTAGTTGTGACATATTAATTTTACAGTGATTAAGTCAAGCAT 888
QY 1016 TCTAGGGTTAATCTAGCTGCATCAGATCATATTCGTCCGGTCTTTTTCGCGCTCAAGTCA 1075
Db 889 GACAAAGTTGACCGG--AATACAGTGAATCCGTGCGGCCCTGACCTGTGAACGAGGTGG 947
QY 1076 TCGCCCAAGCTGGCCCTATCTGGGCATCGGGAGGAGAAAGCCCGTGTCTTTCGCCGCA 1135
Db 948 GGTAGACGGTCTGAACGACGCAAACTGGCGGAACGGTTGGGGTGCAGACGCGCGC 1007
QY 1136 GGTAAAGCGGATGAGAAAGTTTGCAGAGATGACTGCTGTGCAATTAAGTTGACG 1195
Db 1008 TTTACTGGCATTCAGAGAACAGCGGGCGTGTGACGCACTGGCGAAGCCATGTGG 1067
QY 1196 AAAACGACGTTTACATGATGATTCGGGAGAGTGTGGCCATGACGCGCTTAAACGTGA 1255
Db 1068 CGGAAATCATACCTTCGGTGGCGGAGCGAGACGATCGCGCTCATTTCTGATCG 1127
QY 1256 ACTGTTGTTAGGCAACCTGGGATACCAATTCGTGGGCTTTTCGGAACACAGTTCCG 1315
Db 1128 GAA-----TCCGCAAGCTTCAGGCAAGCGCTGTCTCCCTAACCGCAGACAAATGATCT 1181
QY 1316 GATGTCAGCCCGAAGCGCATACGCAACCGAACAATACCGGGAGACGCGGAACCTGCCG 1375
Db 1182 CGAGGAGATCTTCCATCCTACAGTTCTGGCCTGACAGGTCCGGCCGCGACTCTGTAGA 1241
QY 1376 TGCCTGGTGCAGTTAATGACAGCGGTGCGGCTGGGATATTAAGTCAGGAGACG 1435
Db 1242 CGGCTAAGCTTACAGCAATAACCTTGGGCGCTTAAACGGGCTCTGAAGGGGTTTTTGG 1301
QY 1436 GATCTCTGGTGAAGCGGAGAAATGACATGATATACCCGTGATTAACCGCGGGCG 1495
Db 1302 AGCTTCTCGCCCTATAGTGAAGTGTATACAGTTGATTTCTATAGTCACTTAAT 1361
QY 1496 CGCTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGTCACAAT 1555
Db 1362 AGCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGTCACAAT 1421
QY 1556 CCAACAACATACGAGCCGGAAGCATTAAGTGAAGCTGGGGTCTTAATAGTGAAGC 1615
Db 1422 CCAACAACATACGAGCCGGAAGCATTAAGTGAAGCTGGGGTCTTAATAGTGAAGC 1481
QY 1616 TAACTCAGATTAAATTCGTTGCTCACTGCGCGCTTTCAGTGGGAAACCTGTGCGC 1675
Db 1482 TAACTCAGATTAAATTCGTTGCTCACTGCGCGCTTTCAGTGGGAAACCTGTGCGC 1541

QY 1676 CAGCTGCATTATGAATCGGCCAAACGCGCGGAGAGAGCGGTTTGCATATTTGGCGCTCT 1735
Db 1542 CAGCTGCATTATGAATCGGCCAAACGCGCGGAGAGAGCGGTTTGCATATTTGGCGCTCT 1601
QY 1736 TCCGCTTCTCGCTCACTGACTGCTGCGCGCTCGGTGTTGGCTGCGGCGAGCGGTATCA 1795
Db 1602 TCCGCTTCTCGCTCACTGACTGCTGCGCTCGGTGTTGGCTGCGGCGAGCGGTATCA 1661
QY 1796 GCTCACTCAAAAGCGGTAAATACGTTATCCAAGAAATACAGGAGTAACGACGAAAGAAC 1855
Db 1662 GCTCACTCAAAAGCGGTAAATACGTTATCCAAGAAATACAGGAGTAACGACGAAAGAAC 1721
QY 1856 ATGAGCAAAAGGCCAGCAAAAGGCCAGAAACCGTAAAGCCGCGGTGCTGCGTT 1915
Db 1722 ATGAGCAAAAGGCCAGCAAAAGGCCAGAAACCGTAAAGCCGCGGTGCTGCGTT 1781
QY 1916 TTCCATAGGCTCGGCCCTCGTACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGG 1975
Db 1782 TTCCATAGGCTCGGCCCTCGTACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGG 1841
QY 1976 CGAAACCCGACAGGACTAATAAGATACAGGCGTTTCCCTGAAAGCTCCCTGTCGC 2035
Db 1842 CGAAACCCGACAGGACTAATAAGATACAGGCGTTTCCCTGAAAGCTCCCTGTCGC 1901
QY 2036 TCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGGCTTTCCTTCCGGAAGC 2095
Db 1902 TCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGGCTTTCCTTCCGGAAGC 1961
QY 2096 GTGGGCGTTTTCATATGCTCACGCTGATAGTATCTCAATTGGGTAGTGTGCTGC 2155
Db 1962 GTGGGCGTTTTCATATGCTCACGCTGATAGTATCTCAATTGGGTAGTGTGCTGC 2021
QY 2156 AAGCTGGGTGTGTCAAGAACCCCGGTTAGGCCGACCGCTGCGCTTATCCGGTAAAC 2215
Db 2022 AAGCTGGGTGTGTCAAGAACCCCGGTTAGGCCGACCGCTGCGCTTATCCGGTAAAC 2081
QY 2216 TATGCTCTTGAAGTCAACCCGTTAAGACACGACTTATCGCACTGAGCAGCACTGAT 2275
Db 2082 TATGCTCTTGAAGTCAACCCGTTAAGACACGACTTATCGCACTGAGCAGCACTGAT 2141
QY 2276 AACAGATTTCAGAGCGGATAGTAGGGGCTGCTACAGAGTTCTTGAAGTGTGGCT 2335
Db 2142 AACAGATTTCAGAGCGGATAGTAGGGGCTGCTACAGAGTTCTTGAAGTGTGGCT 2201
QY 2336 AACTACGCTACACTAAGAGACATTTGTATCTGCGCTGTGCTGAAGCCAGTTACC 2395
Db 2202 AACTACGCTACACTAAGAGACATTTGTATCTGCGCTGTGCTGAAGCCAGTTACC 2261
QY 2396 TTCCGAAAAAGATGTGTAGCTCTGTATCCGGCAACAAACCAACCGCTGTAGCGGTGT 2455
Db 2262 TTCCGAAAAAGATGTGTAGCTCTGTATCCGGCAACAAACCAACCGCTGTAGCGGTGT 2321
QY 2456 TTTTGTGTTCAGACGACGATTTACGCGCAGAAAAAAGATCTCAAGAAATCTTTG 2515
Db 2322 TTTTGTGTTCAGACGACGATTTACGCGCAGAAAAAAGATCTCAAGAAATCTTTG 2381
QY 2516 ATCTTTTCTACGCGGCTGACGCTGAGTGAACGAAAACTACGTTAAGGATTTGTGTC 2575
Db 2382 ATCTTTTCTACGCGGCTGACGCTGAGTGAACGAAAACTACGTTAAGGATTTGTGTC 2441
QY 2576 ATAGATTTACAAAAAGATCTTCACTAGATCTTTTAAATTAATAAGATTTTAA 2635
Db 2442 ATAGATTTACAAAAAGATCTTCACTAGATCTTTTAAATTAATAAGATTTTAA 2501
QY 2636 TCAATTAAGATTAATAGATTAACCTTGTGTGACAGTTACCAATGCTTAATCAGTAG 2695
Db 2502 TCAATTAAGATTAATAGATTAACCTTGTGTGACAGTTACCAATGCTTAATCAGTAG 2561
QY 2696 GCACTTATTCAGCGATCTGTATTTTGTTCATCATAGTTGCTGATCCCGGTGTG 2755
Db 2562 GCACTTATTCAGCGATCTGTATTTTGTTCATCATAGTTGCTGATCCCGGTGTG 2621
QY 2756 TAGATTAACATGATACGAGAGGCGCTTACATCTGGCCCAAGTGTGATGATACCGCA 2815

Db 2622 TAGATACTAGATACGGAGGGCTTACCATCTGGCCCACTGCTGCAATGATACCGGA 2681
QY 2816 GACCCAGGCTACCGGCTCCGATTTATCAGCAATAAACAGCCAGCCGAAAGGCGAG 2875
Db 2682 GACCCAGGCTACCGGCTCCGATTTATCAGCAATAAACAGCCAGCCGAAAGGCGAG 2741
QY 2876 CGCAGAGTGTGCTCCGCACTTTATCCGCTCCATCCAGTCTATTAATTGTTGGCGGGA 2935
Db 2742 CGCAGAGTGTGCTCCGCACTTTATCCGCTCCATCCAGTCTATTAATTGTTGGCGGGA 2801
QY 2936 GCTAGAGTAAGTATGTCGCACTTAATAGTTGCGCAAGTGTGCACTTGTACAGGC 2995
Db 2802 GCTAGAGTAAGTATGTCGCACTTAATAGTTGCGCAAGTGTGCACTTGTACAGGC 2861
QY 2996 ATCGTGTGTACGCTGCTGTTGGTATGCGCTTCAATCAGCTCCGTTCCCAAGATCA 3055
Db 2862 ATCGTGTGTACGCTGCTGTTGGTATGCGCTTCAATCAGCTCCGTTCCCAAGATCA 2921
QY 3056 AGGCGAGTTACATGATCCCGCATGTTGCAAAAAAGGGTTAGCTCCTTGGTCCG 3115
Db 2922 AGGCGAGTTACATGATCCCGCATGTTGCAAAAAAGGGTTAGCTCCTTGGTCCG 2981
QY 3116 ATCGTGTGCAAGTAAGTGGCCGCAAGTATATCATGATGTTATGGCAGCACTGCAT 3175
Db 2982 ATCGTGTGCAAGTAAGTGGCCGCAAGTATATCATGATGTTATGGCAGCACTGCAT 3041
QY 3176 AATTCTCTTACTGTGATGCCATCCGTAAAGATGCTTTCTGTGACTGTGAGTACTCAAC 3235
Db 3042 AATTCTCTTACTGTGATGCCATCCGTAAAGATGCTTTCTGTGACTGTGAGTACTCAAC 3101
QY 3236 AAGTCATTCTGAGAAATAGTATGCGGAGACCGAGTCTCTTGGCCGCGTCAATACG 3295
Db 3102 AAGTCATTCTGAGAAATAGTATGCGGAGACCGAGTCTCTTGGCCGCGTCAATACG 3161
QY 3296 GATATATCCGCGCACATAGCAGAACTTAAAGTCTCATCATTTGAAAAAGTTCTTCG 3355
Db 3162 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3221
QY 3356 GGGCGAAAACTCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGT 3415
Db 3222 GGGCGAAAACTCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGT 3281
QY 3416 GCACCCCACTGATCTTCAGCATCTTTTACTTTTACACAGCTTTCTGGGTGAGCAAAAA 3475
Db 3282 GCACCCCACTGATCTTCAGCATCTTTTACTTTTACACAGCTTTCTGGGTGAGCAAAAA 3341
QY 3476 GGAAGGCAAAATGCGCAAAAGGAATAAGGCGACACGGAATGTTGAATCTCAT 3535
Db 3342 GGAAGGCAAAATGCGCAAAAGGAATAAGGCGACACGGAATGTTGAATCTCAT 3401
QY 3536 CTCTTCCTTTTTCATATATTATTTGAAGCATTTATCAGGTTATGTTCTCATGAGCGGATAC 3595
Db 3402 CTCTTCCTTTTTCATATATTATTTGAAGCATTTATCAGGTTATGTTCTCATGAGCGGATAC 3461
QY 3596 ATATTGAAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAA 3655
Db 3462 ATATTGAAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAA 3521
QY 3656 GTGCCAC 3662
Db 3522 GTGCCAC 3528

Search completed: January 26, 2005, 03:15:10
Job time : 1771.03 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 08:26:03 ; Search time 1959.66 Seconds
(without alignments)
10737.291 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662
Sequence: 1 ctaaatgtacgcgtataa.....attccccgaagtgcac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 287294193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3662	100.0	3662	US-10-001-189-41	Sequence 41, Appl
2	2665.4	72.8	3637	US-10-211-079-18	Sequence 18, Appl
3	2665.4	72.8	3637	US-10-356-708-32	Sequence 32, Appl
4	2665.4	72.8	3637	US-10-280-913A-32	Sequence 32, Appl
5	2665.4	72.8	3637	US-10-684-134-32	Sequence 32, Appl
6	2665.4	72.8	3637	US-10-637-758-32	Sequence 32, Appl
7	2664.8	72.8	3637	US-10-066-390-4	Sequence 4, Appl1
8	2664.8	72.8	3637	US-10-206-030-4	Sequence 4, Appl1
9	2664.8	72.8	3637	US-10-211-079-4	Sequence 4, Appl1
10	2664.8	72.8	3637	US-10-356-708-17	Sequence 17, Appl
11	2664.8	72.8	3637	US-10-205-772-4	Sequence 4, Appl1
12	2664.8	72.8	3637	US-10-280-913A-17	Sequence 17, Appl

C 13	2664.8	72.8	3637	17	US-10-684-134-17	Sequence 17, Appl
C 14	2664.8	72.8	3637	17	US-10-637-758-17	Sequence 17, Appl
C 15	2656.4	72.5	3637	13	US-10-066-390-3	Sequence 3, Appl1
C 16	2656.4	72.5	3637	13	US-10-206-030-3	Sequence 3, Appl1
C 17	2656.4	72.5	3637	15	US-10-211-079-3	Sequence 3, Appl1
C 18	2656.4	72.5	3637	15	US-10-356-708-31	Sequence 31, Appl
C 19	2656.4	72.5	3637	15	US-10-205-772-3	Sequence 3, Appl1
C 20	2656.4	72.5	3637	17	US-10-280-913A-31	Sequence 31, Appl
C 21	2656.4	72.5	3637	17	US-10-684-134-31	Sequence 31, Appl
C 22	2656.4	72.5	3637	17	US-10-637-758-31	Sequence 31, Appl
C 23	2609.4	71.3	4001	11	US-09-660-763-13	Sequence 13, Appl
C 24	2484.6	67.8	3774	15	US-09-660-763-9	Sequence 9, Appl1
C 25	2469.4	67.4	5277	15	US-10-412-872-16	Sequence 16, Appl
C 26	2464.6	67.3	3908	16	US-10-014-099F-70	Sequence 70, Appl
C 27	2451.8	67.0	3928	15	US-10-128-590-45	Sequence 45, Appl
C 28	2451.8	67.0	3928	15	US-10-128-587A-45	Sequence 45, Appl
C 29	2451.8	67.0	3928	15	US-10-128-578B-45	Sequence 45, Appl
C 30	2446.8	66.8	3715	11	US-09-660-763-11	Sequence 11, Appl
C 31	2446.8	66.8	3927	16	US-10-014-099F-71	Sequence 71, Appl
C 32	2279.2	62.2	4119	9	US-09-993-170-1	Sequence 1, Appl1
C 33	2271.2	62.0	4059	17	US-10-655-702-1	Sequence 1, Appl1
C 34	2230.6	60.9	5250	18	US-10-474-070-51	Sequence 51, Appl
C 35	2216.2	60.5	4205	10	US-09-883-848A-48	Sequence 48, Appl
C 36	2213	60.4	4205	10	US-09-883-848B-35	Sequence 35, Appl
C 37	2172.8	59.3	8167	16	US-10-014-099F-83	Sequence 83, Appl
C 38	2172.8	59.3	8167	16	US-10-014-099F-83	Sequence 83, Appl
C 39	2172.6	59.3	10600	15	US-10-356-708-1	Sequence 1, Appl1
C 40	2172.6	59.3	10600	17	US-10-280-913A-1	Sequence 1, Appl1
C 41	2172.6	59.3	10600	17	US-10-684-134-1	Sequence 1, Appl1
C 42	2172.6	59.3	10600	17	US-10-637-758-1	Sequence 1, Appl1
C 43	2172.6	59.3	10624	15	US-10-356-708-2	Sequence 2, Appl1
C 44	2172.6	59.3	10624	17	US-10-280-913A-2	Sequence 2, Appl1
C 45	2172.6	59.3	10624	17	US-10-684-134-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-10-001-189-41
Sequence 41, Application US/10001189
Publication No. US20020176334A1
GENERAL INFORMATION:
APPLICANT: FRASER JR., MALCOLM J.
APPLICANT: LI, XU
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
FILE REFERENCE: 835910-92098
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,984
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,677
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 3662
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
US-10-001-189-41
Query Match 100.0%; Score 3662; DB 13; Length 3662;
Best local similarity 100.0%; Pred. No. 0;
Matches 3662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
C 1 CTAATGTACGCTATATTTTGTAAATTCGCTAAATTTTGTAAATCAGCTC 60
|||||

Db 1 CTAATGTGTAAGCGTTAATATTTTGTAAATTCGGTTAAATTTTGTAAATCAGCTC 60
QY 61 ATTTTAAACCAATAGAGCCGAAATCGGCAAAATCCCTTATAATCAAAAAGATAGACCGA 120
Db 61 ATTTTAAACCAATAGAGCCGAAATCGGCAAAATCCCTTATAATCAAAAAGATAGACCGA 120
QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGAACAAGAGTCACATTAATTAAGAACTGGAATC 180
Db 121 GATAGGGTTGAGTGTGTTCCAGTTTGAACAAGAGTCACATTAATTAAGAACTGGAATC 180
QY 181 CAACGTCAAAGGCGCAAAACCGTCTATCAGGGGAGATGCGCCACTAAGTAAACCATCACC 240
Db 181 CAACGTCAAAGGCGCAAAACCGTCTATCAGGGGAGATGCGCCACTAAGTAAACCATCACC 240
QY 241 CTAATCAAGTTTGTGGGGTGCAGAGTGCCTTAAGCACTAAATCGAAACCTTAAGAGAG 300
Db 241 CTAATCAAGTTTGTGGGGTGCAGAGTGCCTTAAGCACTAAATCGAAACCTTAAGAGAG 300
QY 301 CCCCCGATTTAGAGCTTGACGGGAAAGCCGGGAAAGCGGAGACGTGCGAAAGAGAGAGAA 360
Db 301 CCCCCGATTTAGAGCTTGACGGGAAAGCCGGGAAAGCGGAGACGTGCGAAAGAGAGAGAA 360
QY 361 AGCGAAAGAGAGCGGCGCTAGAGGCGGTGCAAGTGAAGCGTCAAGCTCGCGTAAACAC 420
Db 361 AGCGAAAGAGAGCGGCGCTAGAGGCGGTGCAAGTGAAGCGTCAAGCTCGCGTAAACAC 420
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCACTTGGCCATTAAGGCTGCG 480
Db 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCACTTGGCCATTAAGGCTGCG 480
QY 481 CAACGTGTGGGAAGGGGAGTCCGTGCGGCTTCTGCTATTAAGCCAGCTGCGCAAGAG 540
Db 481 CAACGTGTGGGAAGGGGAGTCCGTGCGGCTTCTGCTATTAAGCCAGCTGCGCAAGAG 540
QY 541 GGGATGTCTGCAAGGCGATTAAGTGGGTAAAGCGCAGGGTTTCCAGTCAAGACGCTTG 600
Db 541 GGGATGTCTGCAAGGCGATTAAGTGGGTAAAGCGCAGGGTTTCCAGTCAAGACGCTTG 600
QY 601 TAAACGAGCGCGCAGAGAGCGGCGCTCTTCACTTAAGCTTTTGAACCCGTGAGAGAG 660
Db 601 TAAACGAGCGCGCAGAGAGCGGCGCTCTTCACTTAAGCTTTTGAACCCGTGAGAGAG 660
QY 661 GAGAGTCCGGGTGCAAAATGTGTTTACAGCGTATGAGAGATGAAGATGCTCAACAC 720
Db 661 GAGAGTCCGGGTGCAAAATGTGTTTACAGCGTATGAGAGATGAAGATGCTCAACAC 720
QY 721 GCTGAGAAACAGCAGCTAAGATTAACCTAGAAAGATTAATCATATGTCAGTACGTTAA 780
Db 721 GCTGAGAAACAGCAGCTAAGATTAACCTAGAAAGATTAATCATATGTCAGTACGTTAA 780
QY 781 AGATTAATCATGCGTAAATTTGACGATGAGATCTGTAAATGCACTATAGAGCGAAT 840
Db 781 AGATTAATCATGCGTAAATTTGACGATGAGATCTGTAAATGCACTATAGAGCGAAT 840
QY 841 TGAGTACCGGCGCGCGCTGAGGTGCAAGCTATGCAATAGCTTGAATTCGATTCCTGC 900
Db 841 TGAGTACCGGCGCGCGCTGAGGTGCAAGCTATGCAATAGCTTGAATTCGATTCCTGC 900
QY 901 AGCCCGGGGATTCATAGTTCATAGAGCGCGCAGCGGGTGAAGCTCCAGCTTTTGT 960
Db 901 AGCCCGGGGATTCATAGTTCATAGAGCGCGCAGCGGGTGAAGCTCCAGCTTTTGT 960
QY 961 TCCCTTAAAGGAGGTTAATTAATCCATGCGTCAATTTTAAAGCAGATCTTTCTAG 1020
Db 961 TCCCTTAAAGGAGGTTAATTAATCCATGCGTCAATTTTAAAGCAGATCTTTCTAG 1020
QY 1021 GGTAAATCTAGCTGATCAGATCAATATGTCGGGTCTTTTTCGGGTCTGATCGCC 1080
Db 1021 GGTAAATCTAGCTGATCAGATCAATATGTCGGGTCTTTTTCGGGTCTGATCGCC 1080
QY 1081 CAAGCTGCGCTATCTGGGCAATCGGGAGAAAGCCCGTTCCTTTTCCCGGAGGTTG 1140
Db 1081 CAAGCTGCGCTATCTGGGCAATCGGGAGAAAGCCCGTTCCTTTTCCCGGAGGTTG 1140

QY 1141 AAGCGGATGGAAGAGGTTTCCGAGATGACTGCTGTCATATGACGTTAGCGAAAC 1200
Db 1141 AAGCGGATGGAAGAGGTTTCCGAGATGACTGCTGTCATATGACGTTAGCGAAAC 1200
QY 1201 GACGTTTAAACATGATTAATCGGGAAGTGTGCGCATGACGCTTTAACGTTGAATCTG 1260
Db 1201 GACGTTTAAACATGATTAATCGGGAAGTGTGCGCATGACGCTTTAACGTTGAATCTG 1260
QY 1261 TCGTTACGGCCACCTGGGATTAACAGGTTTCTGCGGCTTTTCCGACACAGTTCCGATG 1320
Db 1261 TCGTTACGGCCACCTGGGATTAACAGGTTTCTGCGGCTTTTCCGACACAGTTCCGATG 1320
QY 1321 TCAAGCCGGAAGCCCATCAGAACCCGAAACATTAACGCGGACAGCCGGAACCTGCGCG 1380
Db 1321 TCAAGCCGGAAGCCCATCAGAACCCGAAACATTAACGCGGACAGCCGGAACCTGCGCG 1380
QY 1381 GTGTCAGATTAATGACAGCGGTGCGCGCTGGATTAATGCTACAGCAGAGACGGATATC 1440
Db 1381 GTGTCAGATTAATGACAGCGGTGCGCGCTGGATTAATGCTACAGCAGAGACGGATATC 1440
QY 1441 CTGGCTGATGTCGCGAAGATGAGATGATACCCCGTGAATCCGAGCGGCGCTT 1500
Db 1441 CTGGCTGATGTCGCGAAGATGAGATGATACCCCGTGAATCCGAGCGGCGCTT 1500
QY 1501 GGGCTAATGATGATGATGCTGTTTCTGTGTGAATTTGTATCCGCTCAAAATTCACA 1560
Db 1501 GGGCTAATGATGATGATGCTGTTTCTGTGTGAATTTGTATCCGCTCAAAATTCACA 1560
QY 1561 CAACATGAGAGCGGAAACATTAAGTGAAGCCGTGGGGTCCCTAATGATGATGATCT 1620
Db 1561 CAACATGAGAGCGGAAACATTAAGTGAAGCCGTGGGGTCCCTAATGATGATGATCT 1620
QY 1621 CACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 GCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 GCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 TTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 TTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 CTCAAGGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 CTCAAGGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 AGCAAAAGGCGAGCAAAAGCCGAGAAACCGTAAAGAGCCGCTGAGGCTTTTCA 1920
Db 1861 AGCAAAAGGCGAGCAAAAGCCGAGAAACCGTAAAGAGCCGCTGAGGCTTTTCA 1920
QY 1921 TAGGCTCGCGCGCGCTGAGCATCAAAAGATGAGCTCAAGTCAAGTGGCGAAA 1980
Db 1921 TAGGCTCGCGCGCGCTGAGCATCAAAAGATGAGCTCAAGTCAAGTGGCGAAA 1980
QY 1981 CCGGACAGGATTAATTAAGATGACAGGCTTTTCCCTGGAAGCTCCCTGAGGCTCTC 2040
Db 1981 CCGGACAGGATTAATTAAGATGACAGGCTTTTCCCTGGAAGCTCCCTGAGGCTCTC 2040
QY 2041 TGTTCGAGCCCTGCGCTTAAACGATACCTGCGCTTCTCCCTTGGGAAGCGTGC 2100
Db 2041 TGTTCGAGCCCTGCGCTTAAACGATACCTGCGCTTCTCCCTTGGGAAGCGTGC 2100
QY 2101 GCTTTCTATAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 GCTTTCTATAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
QY 2161 GGGCTGTGTGACAAACCCCGCTTACAGCCGACCCCTGCGCTTATCGGTAATCTATCG 2220
Db 2161 GGGCTGTGTGACAAACCCCGCTTACAGCCGACCCCTGCGCTTATCGGTAATCTATCG 2220

QY	2221	TCCTGAGTCCAA	CCCGGTTAA	CAACAGATTAT	TCGCATCTGGCAGACGCACTGGTTAACG	2280
Db	2221	TCCTGAGTCCAA	CCCGGTTAA	CAACAGATTAT	TCGCATCTGGCAGACGCACTGGTTAACG	2280
QY	2281	GATTAGCAGAGGAG	GTATGTAGGCGG	GGCTACAGAGTTCTTTGAAGTGGTGGCCTA	ACTA	2340
Db	2281	GATTAGCAGAGGAG	GTATGTAGGCGG	GGCTACAGAGTTCTTTGAAGTGGTGGCCTA	ACTA	2340
QY	2341	CGGCTACACT	AGAAAGACAGTAT	TGGTATCTGCGCTCTGCTGAGACCGA	TTACCTTCG	2400
Db	2341	CGGCTACACT	AGAAAGACAGTAT	TGGTATCTGCGCTCTGCTGAGACCGA	TTACCTTCG	2400
QY	2401	AAAAAGAT	TGGTACCTCTT	GTATCGCGGAA	CAAA	2460
Db	2401	AAAAAGAT	TGGTACCTCTT	GTATCGCGGAA	CAAA	2460
QY	2461	TGTTTGCAAGCAG	CAAGTTACGCGCAG	AAAAAAGAGCTC	TAAGAAAGATCCTTTGATCTT	2520
Db	2461	TGTTTGCAAGCAG	CAAGTTACGCGCAG	AAAAAAGAGCTC	TAAGAAAGATCCTTTGATCTT	2520
QY	2521	TTCTACGGGGTCT	GA	CGCTCAGTGGAA	CGAAAACTCAGTTAAGGATTTTGT	2580
Db	2521	TTCTACGGGGTCT	GA	CGCTCAGTGGAA	CGAAAACTCAGTTAAGGATTTTGT	2580
QY	2581	ATTATCAAAAA	AGATCTTCA	CCCTAGATCTCTTTAATTA	ATTAATGAAGTTTAAATCAAT	2640
Db	2581	ATTATCAAAAA	AGATCTTCA	CCCTAGATCTCTTTAATTA	ATTAATGAAGTTTAAATCAAT	2640
QY	2641	CTAAAGTATAT	ATGAGTAA	CTGGTCTGACAGTTACCA	AGTCTTAATCAATGAGGAC	2700
Db	2641	CTAAAGTATAT	ATGAGTAA	CTGGTCTGACAGTTACCA	AGTCTTAATCAATGAGGAC	2700
QY	2701	TATCTCAGCGAT	CTGTCTTATTCGTTCA	TCATAGTTG	CCCTGACCTGCTGAT	2760
Db	2701	TATCTCAGCGAT	CTGTCTTATTCGTTCA	TCATAGTTG	CCCTGACCTGCTGAT	2760
QY	2761	AACTACGATAC	GGGAGGGCTTA	CCATCTGGCCCA	GGTCTCAATGATACCGCAGACCC	2820
Db	2761	AACTACGATAC	GGGAGGGCTTA	CCATCTGGCCCA	GGTCTCAATGATACCGCAGACCC	2820
QY	2821	ACGCTCACGGGCT	CCAGTTATATACGAA	TTAA	CAAGCAGACCGGAGGGCGGACG	2880
Db	2821	ACGCTCACGGGCT	CCAGTTATATACGAA	TTAA	CAAGCAGACCGGAGGGCGGACG	2880
QY	2881	AAGTGGTCTCT	CAACTTTATCCGCTCAT	CACTCTATTA	ATTTGTTGGCGGAAAGCTAG	2940
Db	2881	AAGTGGTCTCT	CAACTTTATCCGCTCAT	CACTCTATTA	ATTTGTTGGCGGAAAGCTAG	2940
QY	2941	AGTAAGTAT	TCGCGAGTTAA	TAGTTTGGCGAA	CGTTGGTTCATTTGTA	3000
Db	2941	AGTAAGTAT	TCGCGAGTTAA	TAGTTTGGCGAA	CGTTGGTTCATTTGTA	3000
QY	3001	GGTGTCAAGCT	CTGTGTTTGGTATG	GGCTTCA	TTCAGCTCCGCTTCCCA	3060
Db	3001	GGTGTCAAGCT	CTGTGTTTGGTATG	GGCTTCA	TTCAGCTCCGCTTCCCA	3060
QY	3061	AGTTACATGAT	ATCCCAATGTTGTG	CAAAAAAGCGGTTAG	CTCTTCGATCGT	3120
Db	3061	AGTTACATGAT	ATCCCAATGTTGTG	CAAAAAAGCGGTTAG	CTCTTCGATCGT	3120
QY	3121	TGTCAAGAA	GTAGTTGGCCG	CAGTGTATCACT	CATGGTTATG	3180
Db	3121	TGTCAAGAA	GTAGTTGGCCG	CAGTGTATCACT	CATGGTTATG	3180
QY	3181	TCCTTACGT	CAATGCGATTC	CGTTAAAGATG	CTTTCTGTGATCTG	3240
Db	3181	TCCTTACGT	CAATGCGATTC	CGTTAAAGATG	CTTTCTGTGATCTG	3240
QY	3241	ATTCTGAGAA	TAGTGTATGCGG	CGACCAAGTTG	CTTCCGCGCTCA	3300
Db	3241	ATTCTGAGAA	TAGTGTATGCGG	CGACCAAGTTG	CTTCCGCGCTCA	3300
QY	3301	TACCGCGC	CAATGACAA	CTTTAAAGTGT	CTATCATTTG	3360
Db	3301	TACCGCGC	CAATGACAA	CTTTAAAGTGT	CTATCATTTG	3360

Db	3301	TACGCGCCCAATGACAACCTTTAAAGTGTCTCATCTTGAAAAAAGTTTTCTGGGGC	3360
QY	3361	AAAATCTCAAGAGATTCTAACGCCGTGTAGATTCAGTGTGAATPACCACCTGTGCACC	3420
Db	3361	AAAATCTCAAGAGATTCTAACGCCGTGTAGATTCAGTGTGAATPACCACCTGTGCACC	3422
QY	3421	CAACTGATCTTCAAGCATCTTTACTCTTCACGAGCGTTTCTGGGTGAGCAAAAACAGAAG	3480
Db	3421	CAACTGATCTTCAAGCATCTTTACTCTTCACGAGCGTTTCTGGGTGAGCAAAAACAGAAG	3482
QY	3481	GCAAAATCCCGCAAAAAAGGGAATPAGGGCGCACAGGAAAGTTGAATCTCATCTCTT	3540
Db	3481	GCAAAATCCCGCAAAAAAGGGAATPAGGGCGCACAGGAAAGTTGAATCTCATCTCTT	3542
QY	3541	CGTTTTCAATATTATTAGACATTTATCAGGGTTATGTCTCATAGACGGATACATTT	3600
Db	3541	CGTTTTCAATATTATTAGACATTTATCAGGGTTATGTCTCATAGACGGATACATTT	3602
QY	3601	TGAATGATTTAGAAAAATAAACAATPAGGGTTCCGCGCATTTCCCCGAAAAAGTCC	3660
Db	3601	TGAATGATTTAGAAAAATAAACAATPAGGGTTCCGCGCATTTCCCCGAAAAAGTCC	3662
QY	3661	AC 3662	
Db	3661	AC 3662	

RESULT 2
US-10-211-079-18/c
; Sequence 18, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
 APPLICANT: Padgett, Hal S.
 APPLICANT: Vaeshongs, Andrew A.
 APPLICANT: Vojdani, Fakraddin S.
 APPLICANT: Smith, Mark L.
 TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
 FILE REFERENCE: P-LG 5381
 CURRENT APPLICATION NUMBER: US/10/211,079
 CURRENT FILING DATE: 2002-08-01
 PRIOR APPLICATION NUMBER: US 10/098,155
 PRIOR FILING DATE: 2002-03-14
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 18
 LENGTH: 3637
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: synthetic construct
US-10-211-079-18

Query Match 72.8%; Score 2665.4; DB 15; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

QY	1	CTAAATGTGAAGCGTTAATATTGTTAAAAATTCGCGTTAAATTTGTTAAATCAGCTC	60
Db	3637	CTAAATGTGAAGCGTTAATATTGTTAAAAATTCGCGTTAAATTTGTTAAATCAGCTC	3578
QY	61	ATTTTAAACCAAATAGGCCGAAATCCGCAAAATCCCTTAATTAACAAAGAAATGACCGA	120
Db	3577	ATTTTAAACCAAATAGGCCGAAATCCGCAAAATCCCTTAATTAACAAAGAAATGACCGA	3518
QY	121	GATAGGGGTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCAATATTAAAGAACTGACCTC	180
Db	3517	GATAGGGGTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCAATATTAAAGAACTGACCTC	3458
QY	181	CAACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACATACCC	240
Db	3457	CAACGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACATACCC	3398

QY 241 CTAATCAAGTTTTTTGGGGTCAGAGTCCCTAAAGCACTAAATCCGAAACCCCTAAAGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGGTCAGAGTCCCTAAAGCACTAAATCCGAAACCCCTAAAGGAG 3338
QY 301 CCCCCGATTTAGACTTGAACGGGAAAGCCGGGAAACCTGCGAGAAAGAAAGGAGAA 360
Db 3337 CCCCCGATTTAGACTTGAACGGGAAAGCCGGGAAACCTGCGAGAAAGAAAGGAGAA 3278
QY 361 AGGAAAGAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGCGGCAACGCTGCGGCGTAAACAC 420
Db 3277 AGGAAAGAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGCGGCAACGCTGCGGCGTAAACAC 3218
QY 421 CACACCCGCGGCTTAATGCGCGCTACAGGGCGCGCTCCATTGCGCATTTCAAGCTCGC 480
Db 3217 CACACCCGCGGCTTAATGCGCGCTACAGGGCGCGCTCCATTGCGCATTTCAAGCTCGC 3158
QY 481 CAACTGTTGGAAAGGCGCATTCGCTGCGGCGCTTCCTGCTATTACGCGACGCTGGCGAAAG 540
Db 3157 CAACTGTTGGAAAGGCGCATTCGCTGCGGCGCTTCCTGCTATTACGCGACGCTGGCGAAAG 3098
QY 541 GGGATGCTCGCAAGCGCATTAAGTTGGGTAACGCCAGGGTTTTTCCCACTCAAGAGCTTG 600
Db 3097 GGGATGCTCGCAAGCGCATTAAGTTGGGTAACGCCAGGGTTTTTCCCACTCAAGAGCTTG 3038
QY 601 TAAAAAGAGCGCGCATGAGCGCGCTCTGCTTCAATTCAGTTTGAACCCGTGGAGAGAGCG 660
Db 3037 TAAAAAGAGCGCGCATGAGCGCGCTCTGCTTCAATTCAGTTTGAACCCGTGGAGAGAGCG 2978
QY 661 GCAAGCTCGCGGTGCAAAATGTTTTTACAGCGTGAAGAGAGAGAGAGAGAGCTGACAC 720
Db 2977 CCGGCGTGGCGCGCGCTCTGAAACTAGTGAATCCCCGGGCTGCGAGAAATCTTAATTTGT 2918
QY 721 GCTGCAAGACAGCGAGCTAGTTAACCTTAAGAAAGATATCATATTGAGAGTACGTTAA 780
Db 2917 ATAGTTATCATGAGCTAGTATGTAATCCAGAGCACTTAACAACTCAAGAAAGACATGT 2858
QY 781 AGATATTC-ATGCGTAAATTTGACGATGAGATCTGTAATACATCACTACTATAGGCGAA 839
Db 2857 GGTACGCTTTTCTGTGGGATCTTTGAAAGGCGAGATTTGTGCAAGGTAATGTGTTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTCAAGGTATCGATTAAGCTTGAATTCGAATTCCTG 899
Db 2797 CTGGTAAAGAGACAGGCGCATTCGCAATTTGAGTAATTTGTGTAATGCTGTGTAAT 2738
QY 900 CAGCCCCGGGATCCATGATTTCTAGAGCGCGCGCAACCGGGTGAAGCTCCAGCTTTTG 959
Db 2737 GAAAGATTCATCTTCAATGTTGTGGGAATTTGAAGTTAGCTTGAATTCATTTCTTTT 2678
QY 960 TTCCCTTTAGAGGGTAAATTAATCCCATGCGCTCAATTTTACGAGACTATCTTTCTA 1019
Db 2677 GTTGTGTCGGTATGTAATGTAATGTAAGTTTGAATTTGTAATGCTGTGTAATGCTG 2618
QY 1020 GGGTTAATCTAGCTGCAATCAGATCATATGCTGGGCTTTTTCGGGCTCAGTATGC 1079
Db 2617 GAATGTTTCAATCTTTTAAATCAATCTTTTAACTCGATC-----GATTAAACAG 2563
QY 1080 CCAAGCTGGCGTATCTGGGCATCGGGAGAGAAAGCCCGTCTTTTCCCGCAGGTT 1139
Db 2562 GGTATCACCTTCAACCTTGACTTCAAGACGCGCTTGTAGTTCCCGTCACTTTGAAAGA 2503
QY 1140 GAAGCGGCAATGAAAGTTTGCAGAGATGACTGCTGCTGATTAAGCTTGAAGCGAA 1199
Db 2502 TATAGTGCCTTCTGTACATTAACCTTGGGCAATGCACTTTGAAAGATATCCGCTT 2443
QY 1200 CGACGTTTACATGATGATTCGGGAAGTGTGCGCATGCAAGCCTTTTAAACGTAATC 1259
Db 2442 CATATGATCCGGTAAACGAGAAACATTTGAACACCATGAGAAATAGTGAACAGTGT 2383
QY 1260 TTGCTCAGGCACTCGGATACAGTTGCTGCGGCTTTTCCGGAACAGTTCCGGAATG 2319
Db 2382 TGCGCATGGAACAGGTAGTTTCCAGTAGTGAATTAATTTAAGTTAGTTCTCCGTA 2323

QY 1320 GTACGCCGAGACCGCATGACAAACCCGAACATACCGGCGACAGCCGGAATCCCGTGC 1379
Db 2322 TGATCATACCTTCAACCTCTCCACTGACAGAAATTTTGTGCCATTTAACTACATTC 2263
QY 1380 GGTGTGCAATTAATGACAGCGGTGGCGGCTGGATATTTAGCTACGAGAGACGGTAT 1439
Db 2262 TAATTCACAGAAAT-----TGGAGCAACTCCAGTGAAGAAAGTTCC 2224
QY 1440 CCGGCGTGAATGCCGAGAAATGACATGATATCCCCGTGAGTTACCCGCGGGCGGCT 1499
Db 2223 TTCTCTTTACTATCGTATCCAGCTTTTGTCTCTTTAAGTGAAGTTAATGCGGCT 2164
QY 1500 TGCGTATCATGATGATAGCTGTTTCCGTGTGAATTTGTTATCCCTCAATTCAC 1559
Db 2163 TGCGTATCATGATGATAGCTGTTTCCGTGTGAATTTGTTATCCCTCAATTCAC 2104
QY 1560 ACAACATACAGACCGGAGCATTAAGTTAAAGCTTGGGTTGCTTAATGATGAGCTTAC 1619
Db 2103 ACAACATACAGACCGGAGCATTAAGTTAAAGCTTGGGTTGCTTAATGATGAGCTTAC 2044
QY 1620 TCACATTAATGCGTTCAGCTGCGCGCTTCCAGTGGGAAACCTGTGTGCGAGC 1679
Db 2043 TCACATTAATGCGTTCAGCTGCGCGCTTCCAGTGGGAAACCTGTGTGCGAGC 1984
QY 1680 TGCAATTAATGATTCGGCCCAACCGCGGGAGAGCGGTTTGCATATTGGCGCTTCCG 1739
Db 1983 TGCAATTAATGATTCGGCCCAACCGCGGGAGAGCGGTTTGCATATTGGCGCTTCCG 1924
QY 1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTGGCGTGGCGGAGGAGTTCAGCTC 1799
Db 1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTGGCGTGGCGGAGGAGTTCAGCTC 1864
QY 1800 ACTCAAGCGGTAATTCGTTATCCACAGATTCAGGGGATTAACGACAGAAAGACATGT 1859
Db 1863 ACTCAAGCGGTAATTCGTTATCCACAGATTCAGGGGATTAACGACAGAAAGACATGT 1804
QY 1860 GAGCAAAAGGCGACAGAAAGCGAGAACGTTAAAGGCGCGTTCGCTGCGTTTTTC 1919
Db 1803 GAGCAAAAGGCGACAGAAAGCGAGAACGTTAAAGGCGCGTTCGCTGCGTTTTTC 1744
QY 1920 ATAGGCTCGCGCCCGCGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1979
Db 1743 ATAGGCTCGCGCCCGCGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAA 1864
QY 1980 ACCGCAAGAGCTATTAAGATACAGGCGTTTCCCGTGAAGTCCCTGTCGCTGC 2039
Db 1683 ACCGCAAGAGCTATTAAGATACAGGCGTTTCCCGTGAAGTCCCTGTCGCTGC 1624
QY 2040 CTGTTCCGACCCCTGACCGGTTACCGGATACCTGTCCGCTTTCTCTTCCGGAAGCGTG 2099
Db 1623 CTGTTCCGACCCCTGACCGGTTACCGGATACCTGTCCGCTTTCTCTTCCGGAAGCGTG 1564
QY 2100 CGCTTTCTATAGCTACGCTGTAGATCTCAGTTGCGTGTAGTGTGCTTCCAGC 2159
Db 1563 CGCTTTCTATAGCTACGCTGTAGATCTCAGTTGCGTGTAGTGTGCTTCCAGC 1504
QY 2160 TTGGCTGTGTCAAGAACCCCGGTTACCGGCGAGCCGCTGCGCTTATCCGTAATC 2219
Db 1503 TTGGCTGTGTCAAGAACCCCGGTTACCGGCGAGCCGCTGCGCTTATCCGTAATC 1444
QY 2220 GTCTTGTGTCAACCGGTAAGACAGACTTATCGCACTGGACAGAGCACTGGTAACA 2279
Db 1443 GTCTTGTGTCAACCGGTAAGACAGACTTATCGCACTGGACAGAGCACTGGTAACA 1384
QY 2280 GGAATTAACAGAGGATATGAGGCGGTCTACAGAGTTCTTGAAGTGTGCTTAATC 2339
Db 1383 GGAATTAACAGAGGATATGAGGCGGTCTACAGAGTTCTTGAAGTGTGCTTAATC 1324
QY 2340 ACCGCTACCTAAGAGACATGATTTGTATCTGCGCTGTGTGAAGCACTTCCG 2399
Db 1323 ACCGCTACCTAAGAGACATGATTTGTATCTGCGCTGTGTGAAGCACTTCCG 1264
QY 2400 GAAAAAGAGTTGTAGCTTGTATCCGCAACAAACACCGCTGTGTAGCGGTGTTTT 2459

Db 1263 GAAAAGAGTGTGTAAGTCTTGTATCCGGCAAAACACCGCTGTGTAAGCGGTGTTT 1204
Qy 2460 TTGTTGGACAGCAGATTAACGGCGAGAAAAAGATCTTCAGAGATCTTGTATCT 2519
Db 1203 TTGTTGGACAGCAGATTAACGGCGAGAAAAAGATCTTCAGAGATCTTGTATCT 1144
Qy 2520 TTTCTACGGGGCTGACCGCTCAGTGGAAAGAAAACTCAGCTTAAGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGGCTGACCGCTCAGTGGAAAGAAAACTCAGCTTAAGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGATCTTCAACCTAGATCTTTTAAATTAATAAGATTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGATCTTCAACCTAGATCTTTTAAATTAATAAGATTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATAGATTAACCTGTTGCTGACAGTTACCAATGCTTAATCAGTGAAGCAC 2699
Db 1023 TCTAAAGTATATAGATTAACCTGTTGCTGACAGTTACCAATGCTTAATCAGTGAAGCAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTTGTTCATCATATGTTGCTGATCTCCCGTCTGTATGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTTGTTCATCATATGTTGCTGATCTCCCGTCTGTATGA 904
Qy 2760 TAACTACGATACGGGAGGCTTAAACCATCTGGCCCGCAGTGTGCAATGATATACGGAGACC 2819
Db 903 TAACTACGATACGGGAGGCTTAAACCATCTGGCCCGCAGTGTGCAATGATATACGGAGACC 844
Qy 2820 CACGCTCACGGGCTTCAGATTTATCAGCAATTAACCAACGACCGGAAAGGGCCGACCA 2879
Db 843 CACGCTCACGGGCTTCAGATTTATCAGCAATTAACCAACGACCGGAAAGGGCCGACCA 784
Qy 2880 GAACTGTCTCTGCACTTTATCCGCTTCATCTCAGTCTATTAATTTGTTGCCGGAAAGCTA 2939
Db 783 GAACTGTCTCTGCACTTTATCCGCTTCATCTCAGTCTATTAATTTGTTGCCGGAAAGCTA 724
Qy 2940 GAGTAAAGTATGTCGCACTTAATAGTTTGGCGAAGTGTGGCATGCTCAAGGCATCG 2999
Db 723 GAGTAAAGTATGTCGCACTTAATAGTTTGGCGAAGTGTGGCATGCTCAAGGCATCG 664
Qy 3000 TGTGTGTCACGCTCGTGTGTTGGTATGAGCTTCAATCAGCTCCGGTTCCTCAAGCATCAGGC 3059
Db 663 TGTGTGTCACGCTCGTGTGTTGGTATGAGCTTCAATCAGCTCCGGTTCCTCAAGCATCAGGC 604
Qy 3060 GAGTTTACATGATCCCCCATGTTTGGCAAAAAAGCGGTTAGCTCTTCGCTCTCCGATCG 3119
Db 603 GAGTTTACATGATCCCCCATGTTTGGCAAAAAAGCGGTTAGCTCTTCGCTCTCCGATCG 544
Qy 3120 TTGTGAGAAATGATGTCGCGCAGAGTATCATCATGAGTTATGAGCAGACATGCAATAT 3179
Db 543 TTGTGAGAAATGATGTCGCGCAGAGTATCATCATGAGTTATGAGCAGACATGCAATAT 484
Qy 3180 CTCTTACTGTATGTCATCCGTAAGATCTTTCTGTGATCAGTGTGATCTCAACCAAGT 3239
Db 483 CTCTTACTGTATGTCATCCGTAAGATCTTTCTGTGATCAGTGTGATCTCAACCAAGT 424
Qy 3240 CATCTGTGAATAGTGTATGCGGCGACCGAAGTTGCTCTTCCCGCGCTCAATAGCGGATA 3299
Db 423 CATCTGTGAATAGTGTATGCGGCGACCGAAGTTGCTCTTCCCGCGCTCAATAGCGGATA 364
Qy 3300 ATACCGCGCAGCATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAAAGTCTTCGCGGGC 3359
Db 363 ATACCGCGCAGCATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAAAGTCTTCGCGGGC 304
Qy 3360 GAAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCACTCGTGAC 3419
Db 303 GAAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCACTCGTGAC 244
Qy 3420 CCAACTGATCTTCAGCATCTTTTACTTTTACCAAGGCTTTTGGGTGAGCAAAAAACAGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTTTACCAAGGCTTTTGGGTGAGCAAAAAACAGAA 184
Qy 3480 GGCAAAATGCGCAAAAAAGGAATTAAGGCGACACGGAATGTTGAATACCTACTACTCT 3539

Db 183 GGCAAAATGCCGCAAAAAAGGAATTAAGGGGACACGGAATGTTGAATCTACTACTCT 124
Qy 3540 TCTTTTTCATATATTAATTTGAAGCATTTATCAGGGTTATTTGTCTATGAGCGGATACATAT 3599
Db 123 TCTTTTTCATATATTAATTTGAAGCATTTATCAGGGTTATTTGTCTATGAGCGGATACATAT 64
Qy 3600 TTGAATGATTTTAAAAAAATTAACAATTAAGGGGTTCCGGCGACATTTCCCGGAAGTGC 3659
Db 63 TTGAATGATTTTAAAAAAATTAACAATTAAGGGGTTCCGGCGACATTTCCCGGAAGTGC 4
Qy 3660 CAC 3662
Db 3 CAC 1

RESULT 3
US-10-356-708-32/c
; Sequence 32, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-356-708-32

Query Match 72.8%; Score 2665.4; DB 15; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 511; Indels 27; Gaps 3;

Qy 1 CTAAATGTAAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTAAATGTAAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
Qy 61 ATTTTAAACCAATAGCGCGGAATCGCAAAATCCCTTAATTAATCAAAAGATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGCGGAATCGCAAAATCCCTTAATTAATCAAAAGATAGACCGA 3518
Qy 121 GATAGGTTGAGTGTTCATGTTTGGAAACAAGTCCACTTAAGAAGCGTGACCTC 180
Db 3517 GATAGGTTGAGTGTTCATGTTTGGAAACAAGTCCACTTAAGAAGCGTGACCTC 3458
Qy 181 CAAGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTAGCTGAACCATGACC 240
Db 3457 CAAGTCAAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTAGCTGAACCATGACC 3398
Qy 3397 CTAAATCAAGTTTGTGGGGTCAAGTCCGTTAAAGCACTAAATCGAAGAACCTTAAGGGAG 3338
Db 301 CCCCCGATTTAAGCTTGAAGCGGGGAAGCGGGGAACGTGTGGCAAGAAAGGAAGGAAGAA 360
Qy 3337 CCCCCGATTTAAGCTTGAAGCGGGGAAGCGGGGAACGTGTGGCAAGAAAGGAAGGAAGAA 3278
Db 361 AGCAAAAGAGCGGGCGCTAGGGGCTGTGGCAAGTGTAGCGGTACGCTGTGCGGTAAACAC 420
Qy 3277 AGCAAAAGAGCGGGCGCTAGGGGCTGTGGCAAGTGTAGCGGTACGCTGTGCGGTAAACAC 3218

QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCCAATGCGCATTCAGGCTCG 480
DB |||||
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCCAATGCGCATTCAGGCTCG 3158
QY 481 CAACTGTTGGAAAGGCGATCGGTGCGGCGCTTTCGCTATTACGCCCACTGGCGAAAG 540
DB 3157 CAACTGTTGGAAAGGCGATCGGTGCGGCGCTTTCGCTATTACGCCCACTGGCGAAAG 3098
QY 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGAGGTTTCCAGTCAGACGTTG 600
DB 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGAGGTTTCCAGTCAGACGTTG 3038
QY 601 TAAAAAGAGCGCCAGTAGAGCGCGCTTCGTTCACTGACGTTTGAACCCGTGAGAGCG 660
DB 3037 TAAAAAGAGCGCCAGTAGAGCGCGCTTCGTTCACTGACGTTTGAACCCGTGAGAGCG 2978
QY 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGGTAGAGAGAGATGAAGATGCGACAC 720
DB 2977 CCGCGGTGCGCGCGCTCTAGAACTAGTGATCCCGCGGCTGAGGAAATTTTATTGTT 2918
QY 721 GCTGCAAAACAGCAGCTAGATTAAACCTAGAAAGATTAATCATTTGTGACGTACGTTA 780
DB 2917 ATAGTTCAATCCATGCGCATGTGTAATCCAGACAGATTAACAATCAAGAGGACCATGT 2858
QY 781 AGATATATC-ATGCGTAAATTTGACGCAATGGATCTGTAAATACGACTCACTATAGGCGCA 839
DB 2857 GGTACGCTTTTCGTTGGGATCTTTTGAAAGGCGAGATTGTGTGACAGGTAAATGTGT 2798
QY 840 TTGGGTACGCGGCGCGCGCTCGAGGTGACGAGTATGATTAACCTGATATGCAATTCCTG 899
DB 2797 CTGGTAAAGAGACAGGCGCATCGCAATGGAGTATTTGTTGATTAATGTCTGCTAGTT 2738
QY 900 CAGCCGCGGAGATCACTAGTTCTAGAGCGCGCGCACCGCGGTGAGACTCAAGCTTTG 959
DB 2737 GAACGATCATCTTAATGTTGTGTGGAAATTTGAAAGTAGCTTGAATTCATTCCTTT 2678
QY 960 TTCCCTTAATGAGGTTAATTAAGATCCCATGCTCAATTTTACGACAGATATCTTTCTA 1019
DB 2677 GTTGTCTGCGGTGATGATTAATGATTTGTTGTTAAATTTGTTCTCGAGTTTGTGTCGA 2618
QY 1020 GGGTATATCTAGCTGATCAGAGATCAATCGTGGGCTTTTTCGCGCTCAAGTATCGC 1079
DB 2617 GAATGTTTCATCTTTTAAATCAATCAATCTTTTAACTCGATC-----GATTACAG 2563
QY 1080 CCAAGCTGCGCTATCTGGGCAATCGGAGAGAAAGACCGGTGCTTTTCCCGAGGTT 1139
DB 2562 GGTATCACTTCAACTTGACTTGACGCGGTGTTGATGTTCCGTCATCTTTGAAAGA 2503
QY 1140 GAAGCGGATGGAAGATTTGCGAGATGACCTGCTGCTGATTAAGCTTGAAGG 1199
DB 2502 TATAGTGGCTTCCTGTATTAATCTTGCGGCACTGCTGAAAGTCAAGTGT 2443
QY 1200 CGCAGCTTACATGATGATTCGGGAGGTGTGGCCATGACGCTTTTACAGTTCGTAAC 1259
DB 2442 CATATGATTCGGATTAAGAGAAAGCATGAACCATGAGAGAAAGTAGTAACAAGTGT 2383
QY 1260 TTGCTTACGCGCACTGGATACAGTTCGTGCGGCTTTTCCGACACAGTTCGGAAG 1319
DB 2382 TGCGCATGGAAGAGTATTTTCAAGTAGTGAATTAATTAATTAAGTTCGTA 2323
QY 1320 GTAGCCCGAAGCGCATCAGCAACCGAAACAATACCGGCGACAGCGCGAATCGCGTGC 1379
DB 2322 TGTAGATCACCTTCAACCTCTCTCACTGACAGAAATTTGTGCCATTAATCAACATC 2263
QY 1380 GGTGTGAGATTAATGACAGCGGTGCGGCTGGGATTAATGTCAGAGGAGAGCGGTAT 1439
DB 2262 TAATTAACAAGAT-----TGGACAATCTCAAGTAAAGTTTC 2224
QY 1440 CCGGTGAGATCGCGAAGATGAATGATGATCCCGGTGAGTTAACCGCGCGCGCT 1499
DB 2223 TTCTCTTACTCATCGATACCACTTTTGTTCCTTTATGAGGTTAATTTGCGCGCT 2164
QY 1500 TGGCGTATCATGATCATAGCTTTCTGTGTAAATTTGTTATCCGCTCAATTCAC 1559

DB 2163 TGGCGTATCATGATCATAGCTGTTTCTGTGTAAATTTGTATCCGCTCAATTCAC 2104
QY 1560 ACACATATCGAGCGGAGACATAAAGTAAAGCTGGGGTGGCTTATAGTAGCTAAC 1619
DB 2103 ACACATATCGAGCGGAGACATAAAGTAAAGCTGGGGTGGCTTATAGTAGCTAAC 2044
QY 1620 TCACATTAATGCGTTCGCTCACTGCGGCTTTCCAGTCGAGTAAACCTGTGTCAGC 1679
DB 2043 TCACATTAATGCGTTCGCTCACTGCGGCTTTCCAGTCGAGTAAACCTGTGTCAGC 1984
QY 1680 TGCATTAATGATTCGCAACGCGCGGAGAGAGCGGTTTGGGTATTTGGCGCTTTCCG 1739
DB 1983 TGCATTAATGATTCGCAACGCGCGGAGAGAGCGGTTTGGGTATTTGGCGCTTTCCG 1924
QY 1740 CTTCTCGCTCACTGACCTGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
DB 1923 CTTCTCGCTCACTGACCTGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864
QY 1800 ACTCAAGCGGTAAATCGGTTATCAGAGATCAGGAGTAAAGCGAAGAAACATGT 1859
DB 1863 ACTCAAGCGGTAAATCGGTTATCAGAGATCAGGAGTAAAGCGAAGAAACATGT 1804
QY 1860 GAGCAAAAGCGGAGAAAGCGAGAAACGTTAAAGCGCGCTTCTGCTGCTTTTCC 1919
DB 1803 GAGCAAAAGCGGAGAAAGCGAGAAACGTTAAAGCGCGCTTCTGCTGCTTTTCC 1744
QY 1920 ATAGCTCGCGCGCTGACAGATCAAAAATGAGGCTCAAGTCAAGGTGCGAA 1979
DB 1743 ATAGCTCGCGCGCTGACAGATCAAAAATGAGGCTCAAGTCAAGGTGCGAA 1684
QY 1980 ACCCGACAGACTTAAGATATCAGAGCGTTTCCCTGGAAGCTTCTGCTGCTGCTC 2039
DB 1683 ACCCGACAGACTTAAGATATCAGAGCGTTTCCCTGGAAGCTTCTGCTGCTGCTC 1624
QY 2040 CTGTTCCGACCTTCCGCTTACCGGATACCTGTCGCTTCTTCTTCCCTTCCGGAACGTTG 2099
DB 1623 CTGTTCCGACCTTCCGCTTACCGGATACCTGTCGCTTCTTCTTCCCTTCCGGAACGTTG 1564
QY 2100 CGCTTCTCACTAGCTCAGCGTGAAGTATCTCAATTCGTTGAGTTCGCTTCAAGC 2159
DB 1563 CGCTTCTCACTAGCTCAGCGTGAAGTATCTCAATTCGTTGAGTTCGCTTCAAGC 1504
QY 2160 TGGGCTGTGTGACGAACCCCGCTTACGCGCGGACCGGCTTATCCGTTAATC 2219
DB 1503 TGGGCTGTGTGACGAACCCCGCTTACGCGCGGACCGGCTTATCCGTTAATC 1444
QY 2220 GTCTTGAATCAACCCGTTAAGACGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2279
DB 1443 GTCTTGAATCAACCCGTTAAGACGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1384
QY 2280 GATTTAGAGAGCGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGTGCTTAATC 2339
DB 1383 GATTTAGAGAGCGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGTGCTTAATC 1324
QY 2340 ACGGCTACATAGAAAGCAGTATTTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2399
DB 1323 ACGGCTACATAGAAAGCAGTATTTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1264
QY 2400 GAAAAAGGTTGTAAGCTTTTGAATCCGCAACAAACCAACCGGTGCTGCTGCTGCTGCT 2459
DB 1263 GAAAAAGGTTGTAAGCTTTTGAATCCGCAACAAACCAACCGGTGCTGCTGCTGCTGCT 1204
QY 2460 TTGTTTCAAGCAGCAGATTAAGCGGAGAAAAAGATCTCAAGAAATCTTTGATCT 2519
DB 1203 TTGTTTCAAGCAGCAGATTAAGCGGAGAAAAAGATCTCAAGAAATCTTTGATCT 1144
QY 2520 TTTTACAGGCTGTGACGCTCAGGTGAAGAAATCTCAGTTAAGGATTTTGTGATGA 2579
DB 1143 TTTTACAGGCTGTGACGCTCAGGTGAAGAAATCTCAGTTAAGGATTTTGTGATGA 1084
QY 2580 GATTATCAAAAAGATCTTCACTGATCTTTTAATTAATTAATGAATTTAATCA 2639

QY 601 TAAACGACGGCACTGAGCGCGCTCTGCTTCATTCAAGTCTTTTGAACCCGTGAGGACGG 660
Db 3037 TAAACGACGGCACTGAGCGCGCTCTGCTTCATTCAAGTCTTTTGAACCCGTGAGGACGG 660
QY 661 GAGAGCTCGCGGTGCAATGTGTTTACAGCGTGAAGAGCAGATGAAATGCTGCACAC 720
Db 2977 CCGCGGTGCGCGCTCTGAGAACTAGTGAATCCCCGGGCTGCGAGAAATCTTATTTGT 2918
QY 721 GCTCAGAAACGAGCTAGATTAACTTAAGAAAGATTAATCAATTTAGAGCTTAA 780
Db 2917 ATAGTTCATCATGATCCCATGTGTATCCAGCAGACATTTACAATCAGAAAGACCATGT 2858
QY 781 AGATTAATC-ATGCGTAAATTTAGCAGATGGAATCTGTAATACACTCTATAGGCGGAA 839
Db 2857 GGTACGCTTTTGTGGGATCTTTGAAAGGGCAGATTTGTGCAAGTAAATGTGTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTGCAAGGTATGATTAAGCTTGAATATGAAATTCCTG 899
Db 2797 CTGGTAAAGAGACAGGGCCATGCGCAATTTGAGATTTTGTGATATGCTGTAGTT 2738
QY 900 CAGCCCGGGGATTCATAGTCTAGAGCGGCGCACCGGGTGAAGCTTCAGCTTTTG 959
Db 2737 GAAAGGATCATCTTCAATGTGTGGGAAATTTGAAGTTAGCTTGAATTCATTTCTTT 2678
QY 960 TTCCCTTAGTAGGGTTAATTAGATCCCATGCGTCAATTTTACGAGACTATCTTTTA 1019
Db 2677 GTTGTGTCGCGTATGTATCATTTGTGTGATTTAAAGTTTGTATGCTGTGTGCTGGA 2618
QY 1020 GGGTAACTTAGCTGATCAGAGATCATATGTCGGGCTTTTTCGGGCTGATGATGCG 1079
Db 2617 GATTTTTCATCTTCTTAAATCAATACCTTTTAACTGATAC-----GATTAACAAG 2563
QY 1080 CCAAGCTGGCGCTATCTGGGATCGGGAGGAAAGCCCGCTTTTCCGCGAGGTT 1139
Db 2562 GGTATCACTTCAAACTTGACTTCAAGACGCGCTTTGATGTTCCCGTATCTTTGAAGA 2503
QY 1140 GAACGGCATGGAAGATTTGCGGAGATGAATGCTGCTGATGATGATGATGATGATG 1199
Db 2502 TATAGTCCGTTCTGATCAATACCTTGGGCAATGCGACTCTTGAAGAAATGATGCGGTT 2443
QY 1200 CGCACTTTACATGATGATTTGCGGAAAGTGTGCGATGCAAGCTTTTAAAGTGAATG 1259
Db 2442 CATATGATCCGATTAACGAAACATTTGAACACCATGAGAGAAATGATGACATGT 2383
QY 1260 TTGCTCAGGCAACCGGATTAACGATTTGCGGCTTTTCCGAGACAGCTTCCGAGT 1319
Db 2382 TGGCATGGAAGATGATTTTCCAGTATGCAATTAATTTAAGTGAAGCTTCCGTA 2323
QY 1320 GTCAAGCCGAAAGCGATGACGAAACCGAACAATTCGCGGACAGCGGAACTGCGGTGC 1379
Db 2322 TGTAGCATCACCTTCACTCTCCACGACAGAAATTTGTGCCATTTAACTACCAATC 2263
QY 1380 GGTGTGAGATTAAAGACAGCGGTGCGCGCTGGATATTAGCTGAGGAGAGCGGTAT 1439
Db 2262 TAATTCAACAGAT-----TGGACAACCTCCAGTGAAGATTC 2224
QY 1440 CCGGCGTGAATGCGCGAAGATGAGATGATATCCCCGGAATTCACCGCGGGCGCGCT 1499
Db 2223 TTTCTCTTATCATCGTATCCAGCTTTTGTTCCTTTAGTGAAGGATTTATTCGCGCT 2164
QY 1500 TGGGTAATCATGATGATGCTGTTTCTGTGTAAATTTGTATCCGCTCAATTTCCAC 1559
Db 2163 TGGCGTAAATCATGATGATGCTGTTTCTGTGTAAATTTGTATCCGCTCAATTTCCAC 2104
QY 1560 ACAACATGAGACCGGAAAGCATTAAGTTAAAGCTGGGGTGGCTAATGATGAGTAAAC 1619
Db 2103 ACAACATGAGACCGGAAAGCATTAAGTTAAAGCTGGGGTGGCTAATGATGAGTAAAC 2044
QY 1620 TCAATTAATTTGCTGCTGATGCTGCTTTCAGTGGGAAACTGTGCTGCGAC 1679
Db 2043 TCAATTAATTTGCTGCTGATGCTGCTTTCAGTGGGAAACTGTGCTGCGAC 1984

QY 1680 TGCATTAATGATTCGCGCAACGCGGGGGAAGAGCGGTTTGCATATTGGCGCTCTTCCG 1739
Db 1983 TGCATTAATGATTCGCGCAACGCGGGGGAAGAGCGGTTTGCATATTGGCGCTCTTCCG 1924
QY 1740 CTTTCTGCTGATCACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1799
Db 1923 CTTTCTGCTGATCACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1864
QY 1800 ACTCAAGCGGTAAATCGTTATTCACAGAAATCAGGGATTAACGAGAAAGACATGT 1859
Db 1863 ACTCAAGCGGTAAATCGTTATTCACAGAAATCAGGGATTAACGAGAAAGACATGT 1804
QY 1860 GAGCAAAAGGCGCAAAAGCCAGAACCTGTAAGAAAGCGCGCTGCTGCTGCTGCTGCTGCTG 1919
Db 1803 GAGCAAAAGGCGCAAAAGCCAGAACCTGTAAGAAAGCGCGCTGCTGCTGCTGCTGCTGCTG 1744
QY 1920 ATAGGCTCGCGCCCTGACGAGCATCAAAATAATGACGCTCAAGTCAAGGTGCGGAA 1979
Db 1743 ATAGGCTCGCGCCCTGACGAGCATCAAAATAATGACGCTCAAGTCAAGGTGCGGAA 1684
QY 1980 ACCGACAGGACTAATTAAGATACAGAGGCTTTCCCTGAAAGCTCCCTGCTGCTGCTGCTG 2039
Db 1683 ACCGACAGGACTAATTAAGATACAGAGGCTTTCCCTGAAAGCTCCCTGCTGCTGCTGCTG 1624
QY 2040 CTTTCCGACCTGCGCTTACCGGATACCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2099
Db 1623 CTTTCCGACCTGCGCTTACCGGATACCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1564
QY 2100 CGCTTCTCATAGCTCAAGCTGATGATCTGATGATCTGATGATGATGATGATGATGATGATG 2159
Db 1563 CGCTTCTCATAGCTCAAGCTGATGATCTGATGATCTGATGATGATGATGATGATGATGATG 1504
QY 2160 TGGGCTGTGACAGAAACCCCGCTTACCGGACCGCTGCGCTTATCCGCTTATCCGCTTATCC 2219
Db 1503 TGGGCTGTGACAGAAACCCCGCTTACCGGACCGCTGCGCTTATCCGCTTATCCGCTTATCC 1444
QY 2220 GTCTTGAATCAACCGGATTAAGACAGACTTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2279
Db 1443 GTCTTGAATCAACCGGATTAAGACAGACTTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1384
QY 2280 GATTAAGAGAGAGATTAATGATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2339
Db 1383 GATTAAGAGAGAGATTAATGATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1324
QY 2340 ACGGCTACCTTAAGAGACAGATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2399
Db 1323 ACGGCTACCTTAAGAGACAGATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1264
QY 2400 GAAAAAGGTTGTGATCTTGAATCCGCAACAAACCAACGCTGCTGCTGCTGCTGCTGCTGCTG 2459
Db 1263 GAAAAAGGTTGTGATCTTGAATCCGCAACAAACCAACGCTGCTGCTGCTGCTGCTGCTGCTG 1204
QY 2460 TTGTTTGAACAGAGATTAACGCGAGAAAAAGATCTCAAGAAAGATCTTTGATCT 2519
Db 1203 TTGTTTGAACAGAGATTAACGCGAGAAAAAGATCTCAAGAAAGATCTTTGATCT 1144
QY 2520 TTTCTACGGGCTGACGCTCAATGGAAGAACTCACTGTTAAGGATTTTGTGATGAT 2579
Db 1143 TTTCTACGGGCTGACGCTCAATGGAAGAACTCACTGTTAAGGATTTTGTGATGAT 1084
QY 2580 GATTATCAAAAGATCTTACCTAGATCTTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2639
Db 1083 GATTATCAAAAGATCTTACCTAGATCTTTTAAATTAATTAATTAATTAATTAATTAATTAAT 1024
QY 2640 TCTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2699
Db 1023 TCTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 964
QY 2700 CTAATCTAGAGATCTGCTGATTTGCTGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2759
Db 963 CTAATCTAGAGATCTGCTGATTTGCTGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 904
QY 2760 TAACTAAGATGAGGAGGCTTACCATCTGCGCCCAAGTGTGCAATGATACCGAGAC 2819

Db 903 TAACTACATACGAGGAGGCTTACCATCTGGCCCAAGTGTCTGCAATGATACCGGACACC 844
Oy 2820 CACGCTCACGCGCTCCAGATTATACAGAAATTAACACAGCAGCCGGAAGGCGGACCA 2879
Db 843 CACGCTCACGCGCTCCAGATTATACAGAAATTAACACAGCAGCCGGAAGGCGGACCA 784
Oy 2880 GAAAGGTCTCTGCACTTTATCCGCTTCATCAGTCTATTATTTGTTGCGGGAAGCTA 2939
Db 783 GAAAGGTCTCTGCACTTTATCCGCTTCATCAGTCTATTATTTGTTGCGGGAAGCTA 724
Oy 2940 GAGTAAGTAGTTCCGCACTTAATGTTTGGCGAAGTGTGGCACTGCTACAGGACATCG 2999
Db 723 GAGTAAGTAGTTCCGCACTTAATGTTTGGCGAAGTGTGGCACTGCTACAGGACATCG 664
Oy 3000 TGGTGTCAAGCTCGTCTGTTGGTATGAGCTTCACTTCCGCTTCCGCAAGCATCAAGC 3059
Db 663 TGGTGTCAAGCTCGTCTGTTGGTATGAGCTTCACTTCCGCTTCCGCAAGCATCAAGC 604
Oy 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGCTCCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGCTCCTCCGATCG 544
Oy 3120 TTGTGAGAAAGTAGTTGGCCGAGGTTATCATCATGAGTTATGAGCAGCATCAATTT 3179
Db 543 TTGTGAGAAAGTAGTTGGCCGAGGTTATCATCATGAGTTATGAGCAGCATCAATTT 484
Oy 3180 CTCTTACTGTCACTGCGCATCCGTAAGATCTTTTCTGTGACGTGAGTACTCAACCAAGT 3239
Db 483 CTCTTACTGTCACTGCGCATCCGTAAGATCTTTTCTGTGACGTGAGTACTCAACCAAGT 424
Oy 3240 CATTTCTGAAATAGTGTATGCGGCGACCGAAGTTGCTCTTGCCCGGCTCAATACGGATA 3299
Db 423 CATTTCTGAAATAGTGTATGCGGCGACCGAAGTTGCTCTTGCCCGGCTCAATACGGATA 364
Oy 3300 ATACCGGCGCATATGACGAACTTTAAAGTCTCATCATGAGGAAAGGTTCTTCGGGAGC 3359
Db 363 ATACCGGCGCATATGACGAACTTTAAAGTCTCATCATGAGGAAAGGTTCTTCGGGAGC 304
Oy 3360 GAAAACTCTCAAGGATCTTACCGCTGTGTAGATTCAGATTCAGATTCAGATTCAGATTC 3419
Db 303 GAAAACTCTCAAGGATCTTACCGCTGTGTAGATTCAGATTCAGATTCAGATTCAGATTC 244
Oy 3420 CCAACTGATCTTCAGCATCTTTTACTTCAACGAGGTTTCTGGGTGAGCAAAACAGGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTCAACGAGGTTTCTGGGTGAGCAAAACAGGAA 184
Oy 3480 GGGAAAAATGCGGCAAAAAAGGAAATAGGCGGCAACGGAATGTTGAATATCTCATCTCT 3539
Db 183 GGGAAAAATGCGGCAAAAAAGGAAATAGGCGGCAACGGAATGTTGAATATCTCATCTCT 124
Oy 3540 TCCCTTTTCAATATATTATTAAGCATTTTATCAGGAGTTATGTTCTCATGAGCGGATCATAT 3599
Db 123 TCCCTTTTCAATATATTATTAAGCATTTTATCAGGAGTTATGTTCTCATGAGCGGATCATAT 64
Oy 3600 TTGAATGATTTTGAATAAATAAACAATAGGGGTTCCGCGCAATTTCCCGAAAAAGTGC 3659
Db 63 TTGAATGATTTTGAATAAATAAACAATAGGGGTTCCGCGCAATTTCCCGAAAAAGTGC 4
Oy 3660 CAC 3662
Db 3 CAC 1

RESULT 5
US-10-684-134-32/c
; Sequence 32, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5192-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134

Query Match 72.8%; Score 2665.4; DB 17; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

1 CTAATGTAAGCGTTAATATTTGTTAAATGCGGTTAAATTTGTTAAATCAGCTC 60
3637 CTAATGTAAGCGTTAATATTTGTTAAATGCGGTTAAATTTGTTAAATCAGCTC 3578
Oy 61 ATTTTAAACCAATAGCGGCAATCGCAAAATCCCTTAATCAAAAGAAATAGACCGA 120
Db 3577 ATTTTAAACCAATAGCGGCAATCGGCAAAATCCCTTAATCAAAAGAAATAGACCGA 3518
Oy 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATTCACATTTAAAGAACTGAGCTC 180
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATTCACATTTAAAGAACTGAGCTC 3458
Oy 181 CAAGTCAGAGGAGGAAAAACCGTCTATCAGGGGAGTGGCCCATAGTGAAACATACACC 240
Db 3457 CAAGTCAGAGGAGGAAAAACCGTCTATCAGGGGAGTGGCCCATAGTGAAACATACACC 3398
Oy 241 CTAATCAAGTTTGTGGGGTTCAGAGTCCGCTAAAGCACTAAATCGGAACCTTAAGGAG 300
Db 3397 CTAATCAAGTTTGTGGGGTTCAGAGTCCGCTAAAGCACTAAATCGGAACCTTAAGGAG 3338
Oy 301 CCCCCGATTTAGACTTGAACGGGAAAGCGGCGAAGCTGCGAAGAAAGGAAAGAA 360
Db 3337 CCCCCGATTTAGACTTGAACGGGAAAGCGGCGAAGCTGCGAAGAAAGGAAAGAA 3278
Oy 361 AGCGAAGAGAGCGGCGCTAGGGGCTGAGGAGTGAAGTGAAGCGGTCAAGCGTGCAGTAAACAC 420
Db 3277 AGCGAAGAGAGCGGCGCTAGGGGCTGAGGAGTGAAGTGAAGCGGTCAAGCGTGCAGTAAACAC 3218
Oy 421 CACACCGCGCGGCTTAATGCGCGCTAACAAGGCGGCTCCATTCGSCATTCAGAGCTGCG 480
Db 3217 CACACCGCGCGGCTTAATGCGCGCTAACAAGGCGGCTCCATTCGSCATTCAGAGCTGCG 3158
Oy 481 CAACTGTTGGAGAGGCGATGAGTGGCGGCTCTTGGCTATTACGCCAGCTGCGGAAAG 540
Db 3157 CAACTGTTGGAGAGGCGATGAGTGGCGGCTCTTGGCTATTACGCCAGCTGCGGAAAG 3098
Oy 541 GGGATGTGCGCAAGGCAATTAATGAGTGAAGCGGAGGTTTCCAGTCAAGAGTGG 600
Db 3097 GGGATGTGCGCAAGGCAATTAATGAGTGAAGCGGAGGTTTCCAGTCAAGAGTGG 3038
Oy 601 TAAACGACGCGCAGTAGAGCGCGCTGTTCAATCAAGTTTGAACCCGAGGAGAGCG 660
Db 3037 TAAACGACGCGCAGTAGAGCGCGCTGTTCAATCAAGTTTGAACCCGAGGAGAGTGC 2978
Oy 661 GCAAGCTCGCGGTGCAAAATGTTTAAACAGCTGATGAGAGCAATGATGCTGACAC 720
Db 2977 CCGGCGGTGCGCGCTGCTAGAACTAGTGAATCCCGGCGCTGACGAGAAATCTTATTTGT 2918
Oy 721 GCTCAGAACGACGAGCTAGATTAAACCTGAAAGATTAATCATTTGTGACGTGCTTAA 780

Db 2917 ATAGTTCATCATCCATCCATGCTATCCACGACAGATTACAAACTCAAGAAAGACCATGT 2858
QY 781 AGATTAATC-ATGGGTAAATTTGACGCATGGGATCTGTAAATACACTCATATAGGCGAA 839
Db 2857 GGTACGCTTTTGGTGGGATCTTTGAAAGGCGAGATTTGTGTGCAACAGTTAAAGTGTGT 2798
QY 840 TTGGGTACCGGGCCCCCTCGAGGTGCAAGCTATGCAATTAAGCTTGATATGAAATTCCTG 899
Db 2797 CTGGTAAAGGACAGGCGCATCGCAATTTGGAGTATTTTGTGTAATGTCTGTAGT 2738
QY 900 CAGCCCGGGGATCCACTAGTTCTAGACCGCGCCACCGCGGTGAGCTCCAGCTTTTG 959
Db 2737 GAAAGCATCATTTCAATGTGTGGCAATTTTGAAGTTAGCTTTGATTCATCTTTT 2678
QY 960 TTCCCTTATGAGGGTTAATTAGATCCCATGCGCTCAATTTTACGACAGATCTTTCTA 1019
Db 2677 GTTGTCTGCGGTGATGTAATGTAATGTTGTGATTAAAGTTGTACTCAAGTTGTGTCCGA 2618
QY 1020 GGGTAAATCTAGCTGCATCAGGATCAATACGTGGGATCTTTTTCGGGCTCAATCCG 1079
Db 2617 GAATGTTTCATCTTCTTTAAATCAATACCTTTTATCTGATAC-----GATTAACAAG 2563
QY 1080 CCAAGCTGGCGCTTATCGGGCATGGGGAGAAAGAACCCGTCCTTTTCCGCGAGATT 1139
Db 2562 GGTATCACTTCAAACTTGACTTGACAGCGCTTGTGTAGTTCCGTCATCTTTGAAAGA 2503
QY 1140 GAACGGCATGAAAGAGTTTGCAGAGATGATGCTGCGTGAATTGACGTTAGCGGAAA 1199
Db 2502 TATAGTCCGTCTCTGTATCATACCTTCGGGCATGGGACTTGAAGAAAGTCATGCGTTT 2443
QY 1200 CGCAGTTTACCATGATGATTTGGGAAAGGTGTGCGCATGACAGCTTTTAAACGTTGAACTG 1259
Db 2442 CATATGATCGGATTAACGAGAAAGCATTTGAACACCATGAGAAAGATGACAAAGTGT 2383
QY 1260 TTGGTTCAGCCACTGGGATTAACAGTTGTGCGGCTTTTCGGAACAAGTCCGAGTG 1319
Db 2382 TGCCCATGGAACAGTATGTTTTCAGTATGTGCAAAATTAATTAAATGTTAACTTTCCGTA 2323
QY 1320 GTACGCCGAAGCGCATGACGAACCCGAACATACCGCGCAACAGCGGAACTGCGCTGCC 1379
Db 2322 TGTAGCATCACTTCACTTCCACTGACAGAAATTTGTGCCATTAATCATACCATC 2263
QY 1380 GGTGTCAATTAATGACAGCGGTGCGGCTGGGATATTAAGTCAAGAGAGCGGTAT 1439
Db 2262 TAATTCACAGAAAT-----TGGGCAACTCAGTGAAGAAATTC 2224
QY 1440 CCTGGCTGATGTCGCGAGAAATGACATGATACCCCGTGAAGTTACCCGCGGCGCGCT 1499
Db 2223 TTCTCTTTACTCATCGGTACCCAGCTTTGTTCCTTTAGTGAAGGTTAATTGCCGCT 2164
QY 1500 TGGCGTATCATGTGATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTCAC 1559
Db 2163 TGGGTAAATCATGTGATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTCAC 2104
QY 1560 ACAACATACGAGCGGGAAGCAATAAGTAAAGCTTGGGTGCTAATGATGAGCTAAC 1619
Db 2103 ACAACATACGAGCGGGAAGCAATAAGTAAAGCTTGGGTGCTAATGATGAGCTAAC 2044
QY 1630 TCAATTAATTTGGCTGCTCACTGCGCTTTTCCAGTCCGGAAACCTGTGCTGACG 1679
Db 2043 TCAATTAATTTGGCTGCTCACTGCGCTTTTCCAGTCCGGAAACCTGTGCTGACG 1684
QY 1680 TGCATTAATGATGCGGCAACGCGCGGGAAGGCGGTTTGGGTATTTGGGCGCTTTCCG 1739
Db 1983 TGCATTAATGATGCGGCAACGCGCGGGAAGGCGGTTTGGGTATTTGGGCGCTTTCCG 1924
QY 1740 CTTCCTGCTCACTGACTGCTGCTGCGCTCGGTGCTGCGTGGGAGACCGGTATCAGCTC 1799
Db 1923 CTTCCTGCTCACTGACTGCTGCTGCGCTCGGTGCTGCGTGGGAGACCGGTATCAGCTC 1864
QY 1800 ACTCAAGGCGGTAATACGTTATTCACAAGATCAGGGGATTAACGAGAAAGAACTGT 1859
Db 1863 ACTCAAGGCGGTAATACGTTATTCACAAGATCAGGGGATTAACGAGAAAGAACTGT 1804

QY 1860 GAGCAAAAGCCACGAAAGCCGAGAACCGTAAAGAGCCGCGTGTGCTGCTTTTCC 1919
Db 1803 GAGCAAAAGCCACGAAAGCCGAGAACCGTAAAGAGCCGCGTGTGCTGCTTTTCC 1744
QY 1920 ATAGGCTCGGCCCCCTGACAGAGCATCACAAAATATGAGCTCAATCTCAGAGTGGCGAA 1979
Db 1743 ATAGGCTCGGCCCCCTGACAGAGCATCACAAAATATGAGCTCAATCTCAGAGTGGCGAA 1684
QY 1980 ACCGACAGAGCTAATAAGATTAACAGGCGTTTCCCCGAAAGCTCCCTGTGCGCTTC 2039
Db 1683 ACCGACAGAGCTAATAAGATTAACAGGCGTTTCCCCGAAAGCTCCCTGTGCGCTTC 1624
QY 2040 CTGTTCCGACCTTGCCGCTTACCGGATACCTGTGCGGCTTTCCCTTCCGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCTTGCCGCTTACCGGATACCTGTGCGGCTTTCCCTTCCGGAAGCGTGG 1564
QY 2100 CGCTTTCTCATACCTCAAGCTGTAGGTATCTCAAGTTCGAGTGTGAGTGTGCTGCTCAAGC 2159
Db 1563 CGCTTTCTCATACCTCAAGCTGTAGGTATCTCAAGTTCGAGTGTGAGTGTGCTGCTCAAGC 1504
QY 2160 TGAGCTGTGTGCAAGAACCCCGTTCAGCGCCGACCGCTGCGGCTTATCCGTTAACTATC 2219
Db 1503 TGGGCTGTGTGCAAGAACCCCGTTCAGCGCCGACCGCTGCGGCTTATCCGTTAACTATC 1444
QY 2220 GTCTGAGTCCAAACCGGTTAAGACAGACTTATGCGCATGCGCAGACCACTGTAACA 2279
Db 1443 GTCTGAGTCCAAACCGGTTAAGACAGACTTATGCGCATGCGCAGACCACTGTAACA 1384
QY 2280 GGAATTAAGACAGAGATGATGAGCGGTGCTTACAGAGTTCTTGAAGTGTGCTTAACT 2339
Db 1383 GGAATTAAGACAGAGATGATGAGCGGTGCTTACAGAGTTCTTGAAGTGTGCTTAACT 1324
QY 2340 ACGGCTACACTAAGAAAGACATATTTGTATCTGCGCTGCTGTAAGCCAGTTACTG 2399
Db 1323 ACGGCTACACTAAGAAAGACATATTTGTATCTGCGCTGCTGTAAGCCAGTTACTG 1264
QY 2400 GAAAAAGATTGTAGTCTTGTATCCGGCAACAAACCCGCTGTAGCGGTGTTTTT 2459
Db 1263 GAAAAAGATTGTAGTCTTGTATCCGGCAACAAACCCGCTGTAGCGGTGTTTTT 1204
QY 2460 TTGTTTCAAGACAGACATTAACGCGCAAAAAAGATCTCAAGAAAGATCTTTGATCT 2519
Db 1203 TTGTTTCAAGACAGACATTAACGCGCAAAAAAGATCTCAAGAAAGATCTTTGATCT 1144
QY 2520 TTCTACGGGGTCTGACGCTCACTGGAAGAAACCTCAAGTTAAGGATTTTGTCTATGA 2579
Db 1143 TTCTACGGGGTCTGACGCTCACTGGAAGAAACCTCAAGTTAAGGATTTTGTCTATGA 1084
QY 2580 GATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAAGATGAAGTTTAAATCAA 2639
Db 1083 GATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAAGATGAAGTTTAAATCAA 1024
QY 2640 TCTAAAGTATATATAGTAACTTGTGTGACAGTTACCAATGCTTAAATCACTAGAGCAC 2699
Db 1023 TCTAAAGTATATATAGTAACTTGTGTGACAGTTACCAATGCTTAAATCACTAGAGCAC 964
QY 2700 CTATCTCAAGCATCTGTCTATTTTCTTATCATAGTTGCTGACTCCCGGTGTGAGA 2759
Db 963 CTATCTCAAGCATCTGTCTATTTTCTTATCATAGTTGCTGACTCCCGGTGTGAGA 904
QY 2760 TAACTACGATACGAGAGGCTTAAACATGCGCCCACTGCTGCAATGATACCGAGAC 2819
Db 903 TAACTACGATACGAGAGGCTTAAACATGCGCCCACTGCTGCAATGATACCGAGAC 844
QY 2820 CAGGCTACCGGCTTCAGATTTATCAGCAATAAACAGCAGCGGAAAGGCGAGCGCA 2879
Db 843 CAGGCTACCGGCTTCAGATTTATCAGCAATAAACAGCAGCGGAAAGGCGAGCGCA 784
QY 2880 GAAGTGTCTCGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTTTCGCGGAAGCTA 2939
Db 783 GAAGTGTCTCGCAACTTTTATCCGCTCCATCCAGTCTATTAATTTTTCGCGGAAGCTA 724

Qy	2940	AAGTAAAGAGTATCCGCACATTAATAGTTTGGCCAAAGTTGTTGCCATTGCTACAGGCATCG	2939
Db	723	GAGTAAGAGTATCCGCACATTAATAGTTTGGCCAAAGTTGTTGCCATTGCTACAGGCATCG	664
Qy	3000	TGATGTCAAGCTCCGCTGTTTGGTATAGCTTCAATTCAGCTCCGGTATCCCAAGATCAAGGC	3059
Db	663	TGATGTCAAGCTCCGCTGTTTGGTATAGCTTCAATTCAGCTCCGGTATCCCAAGATCAAGGC	604
Qy	3060	GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCTCCGATCG	3119
Db	603	GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCTCCGATCG	544
Qy	3120	TTGTACAGAAGTAAGTTGGCCGCGAGTGTATCACTCATAGTTATAGCAGCATCTCATATT	3179
Db	543	TTGTACAGAAGTAAGTTGGCCGCGAGTGTATCACTCATAGTTATAGCAGCATCTCATATT	484
Qy	3180	CTCTTACTGTATGTCATCGGTAAGATGCTTTCTGTGACGTGGTAGTACTCAACCAAGT	3239
Db	483	CTCTTACTGTATGTCATCGGTAAGATGCTTTCTGTGACGTGGTAGTACTCAACCAAGT	424
Qy	3240	CATTCTGAAATAGTGTATGCGGCGACCGAATGCTCTTGCCCGCGCTCAATACGGGATA	3299
Db	423	CATTCTGAAATAGTGTATGCGGCGACCGAATGCTCTTGCCCGCGCTCAATACGGGATA	364
Qy	3300	ATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGAAAAAGCTTCTCGGGGC	3359
Db	363	ATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATTTGAAAAAGCTTCTCGGGGC	304
Qy	3360	GAAAACTCTCAAGATCTTACCGCTGTGGAATCCAGTTCGATTAACCCACTGTCAC	3419
Db	303	GAAAACTCTCAAGATCTTACCGCTGTGGAATCCAGTTCGATTAACCCACTGTCAC	244
Qy	3420	CCAATGATCTTCAGCATCTTTTACTTTCACACAGCTTTCTGGGTGACAAAAACAGAA	3479
Db	243	CCAATGATCTTCAGCATCTTTTACTTTCACACAGCTTTCTGGGTGACAAAAACAGAA	184
Qy	3480	GGCAAAATGCGCGCAAAAAAGGAAATTAAGGGCGACACGGAAATGTGAATACTACTCT	3539
Db	183	GGCAAAATGCGCGCAAAAAAGGAAATTAAGGGCGACACGGAAATGTGAATACTACTCT	124
Qy	3540	TCCTTTTTCATATTATTTGAGCATTTATCAGGGTATATGTCATATGACGGATATCATAT	3599
Db	123	TCCTTTTTCATATTATTTGAGCATTTATCAGGGTATATGTCATATGACGGATATCATAT	64
Qy	3600	TTGAATGATTTAGAAAAATAAACAAATAGGGGTTCCGCGACATTTCCCGAAAGATGC	3659
Db	63	TTGAATGATTTAGAAAAATAAACAAATAGGGGTTCCGCGACATTTCCCGAAAGATGC	4
Qy	3660	CAC 3662	
Db	3	CAC 1	
RESULT 6			
US-10-637-758-32/c			
; Sequence 32, Application US/10637758			
; Publication No. US20040180352A1			
GENERAL INFORMATION:			
APPLICANT: Large Scale Biology Corporation			
TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS			
FILE REFERENCE: LSBC-PI65392-CIP			
CURRENT APPLICATION NUMBER: US/10/637,758			
CURRENT FILING DATE: 2003-08-08			
PRIOR APPLICATION NUMBER: US/10/280,913			
PRIOR FILING DATE: 2002-10-25			
PRIOR APPLICATION NUMBER: 60/402,342			
PRIOR FILING DATE: 2002-08-08			
PRIOR APPLICATION NUMBER: 10/066,390			
PRIOR FILING DATE: 2002-02-01			
PRIOR APPLICATION NUMBER: 60/266,785			
PRIOR FILING DATE: 2001-02-14			
PRIOR APPLICATION NUMBER: 60/266,386			
PRIOR FILING DATE: 2002-02-02			

Query Match	Beef Local Similarity	83.7%	Score 2665.4;	DB 17;	Length 3637;
Matches 3065; Conservative	0;	Mismatches 571;	Indels 27;	Gaps 3;	
/ NUMBER OF SEQ ID NOS: 35 / SOFTWARE: Patentin version 3.1 / SEQ ID NO 32 / LENGTH: 3637 / TYPE: DNA / ORGANISM: Artificial Sequence / FEATURE: / PEATIRE: / OTHER INFORMATION: Encodes Cycle 3 BFP gene US-10-637-758-32					
Query	1	CTAAATTGTAAGCGTTATATTTTGT	AAATTCGCGTTAAATTTTGT	AAATTCAGCTC	60
Db	3637	CTAAATTGTAAGCGTTATATTTTGT	TAAATTCGCGTTAAATTTTGT	AAATTCAGCTC	3578
QY	61	ATTTTAAACCAATAGCGCGAAATCG	CAAAATCCCTTAATCAAAAGATAG	CCGA 120	
Db	3577	ATTTTAAACCAATAGCGCGAAATCG	CAAAATCCCTTAATCAAAAGATAG	CCGA 3518	
QY	121	GATAGGGTGTAGTGTGTTCAGTTTG	GAATAAGTCACTATTAAGAACTG	AGATC 180	
Db	3517	GATAGGGTGTAGTGTGTTCAGTTTG	GAATAAGTCACTATTAAGAACTG	AGATC 3458	
QY	181	CAACGTCAAAAGGCGAAACCGTCTA	TACAGGGGATGAGCCCACTGAA	ACATCAC 240	
Db	3457	CAACGTCAAAAGGCGAAACCGTCTA	TACAGGGGATGAGCCCACTGAA	ACATCAC 3398	
QY	241	CTAATCAAGTTTTTGGGGTGAAGTG	CCGTAAAGCACTAATTCGAAAC	CTTAAAGGAG 300	
Db	3397	CTAATCAAGTTTTTGGGGTGAAGTG	CCGTAAAGCACTAATTCGAAAC	CTTAAAGGAG 3338	
QY	301	CCCCCGATTAGACTTGAACGGGAA	ACCGCGCAAGTGGCGAAGAAAG	AGAGAA 360	
Db	3337	CCCCCGATTAGACTTGAACGGGAA	ACCGCGCAAGTGGCGAAGAAAG	AGAGAA 3278	
QY	361	AGCGAAAGGACCGGGCGCTAGAGG	CGTGAAGTGAAGCGGTCAAGCT	CGCGCTAAC 420	
Db	3277	AGCGAAAGGACCGGGCGCTAGAGG	CGTGAAGTGAAGCGGTCAAGCT	CGCGCTAAC 3218	
QY	421	CACACCCGCGCGTTATATGCGCG	CGCTACAGGGCGCGTCCATTG	CGCATTCAG 480	
Db	3217	CACACCCGCGCGTTATATGCGCG	CGCTACAGGGCGCGTCCATTG	CGCATTCAG 3158	
QY	481	CAACTGTTGGGAAGGGCGATGCTG	TGCGGGCTCTTCTCGTATTAAC	CGCAGTGG 540	
Db	3157	CAACTGTTGGGAAGGGCGATGCTG	TGCGGGCTCTTCTCGTATTAAC	CGCAGTGG 3098	
QY	541	GGGATGTCGTCGAAGGCGATTAA	TGTTGGGTAAAGCCAGGGTTT	CCAGTACAG 600	
Db	3097	GGGATGTCGTCGAAGGCGATTAA	TGTTGGGTAAAGCCAGGGTTT	CCAGTACAG 3038	
QY	601	TAAACGACGCGCAGTGAACGCG	CGCTGTTCAATCAAGTTTTG	AACCCGAGAG 660	
Db	3037	TAAACGACGCGCAGTGAACGCG	CGCTGTTCAATCAAGTTTTG	AACCCGAGAG 2978	
QY	661	GCAGACTGCGCGTCAAAATGTG	TTTACAGCGTATGAGAGCAG	TGAAGATG 720	
Db	2977	GCAGACTGCGCGTCAAAATGTG	TTTACAGCGTATGAGAGCAG	TGAAGATG 2918	
QY	721	GCTGCAAGAACGACGATAGATT	AACCTTAGAAAGATATATATT	TGTGACGT 780	
Db	2917	GCTGCAAGAACGACGATAGATT	AACCTTAGAAAGATATATATT	TGTGACGT 2858	
QY	781	AGATATATC-ATGCGTAAATTA	TGAGCATGGGATCTGAATAC	AGACTAC 839	
Db	2857	AGATATATC-ATGCGTAAATTA	TGAGCATGGGATCTGAATAC	AGACTAC 2798	
QY	840	TTGGGTACCGGGCCCCCTCGA	AGTGCAGCGTATGACTTGA	TATTCGATT 899	
Db	2797	TTGGGTACCGGGCCCCCTCGA	AGTGCAGCGTATGACTTGA	TATTCGATT 2718	

QY 900 CAGCCGGGGGATCCACTGATTCTTGAAGCGCGCCGACCGCGGTGGAGCTCCAGCTTTTG 959
DB 2737 GAGGGATCCATCTTCAATGTTGTGGCAATTTTGAAGTTAGTTGATTCATTTCTTT 2678
QY 960 TTCCTTTTGAAGGGTTAATTAATAGATCCCATGGGTCAATTTTAAAGCAATATCTTTCTA 1019
DB 2677 GTTGTCTGCCGTGATGTATACATTTGTGAGTTAAAGTTGATCTGAGTTTGTGTCCGA 2618
QY 1020 GGGTTAATTAAGCTGCATCAGGATCAATATCGTGGGTCTTTTTCGGGCTGAGTATCCG 1079
DB 2617 GAATGTTTCCATCTCTTTAAATCAATCTTTTAAGTCAATAC-----GATTAACAAG 2563
QY 1080 CCAAGCTGGCGCTATCTGGGCAATCGGGAGGAAGAACCCGTGCTTTTCCCGCAGGTT 1139
DB 2562 GGTATCACTTCAAACTTGACTTCAGCACGCTCTTGAATTCCTGATCTTTGAAAGA 2503
QY 1140 GAAGCGGAGGAAAGATTTGCGAGAGATGACTGTGCTGATTCATTCGTTGAGCGAAA 1199
DB 2502 TATAGTGGCTTCTGTATCAATACCTTCGGGCATGGCACTGTGAAAAAGTCATGCGGTTT 2443
QY 1200 CGAGCTTTACCATGATGATTCGGGAAGTGTGGCAATGCAAGCTTTTAAAGTGAATCTG 1259
DB 2442 CATATGATCCGGATTAACGAAAAAGCATTTGAACCACTGAGAAAGATAGTGAACAAGTG 2383
QY 1260 TTCTGTTAGGCACTGGGATACCAAGTTCTGTGGGCTTTTCCGACACAGTTCCGGATG 1319
DB 2382 TGGCCATGGAACAGTATGTTTCCAGATGCAAAATTAATTAAGTAAAGCTTTCCGTA 2323
QY 1320 GTACGCCGAAAGCCGATCAGCAACCCGAAACATACCGGAGACAGCCGAACTGCCGTCC 1379
DB 2332 TGTAGCAATCACTTCACTCTCCAGTACAGAAAAATTTGTGCCATTAACATCAACATC 2263
QY 1380 GGTGTGAGATTATGACACGGGTGGCGGCTGGGATATTACGTACAGAGAGAGGGAT 1439
DB 2262 TAATTCAACAAGAT-----TGGACAATCTCAAGTGAAGTTTC 2224
QY 1440 CTGTGCTGATGTCGGAGAAATGACATGATATCCCGTGAATACCGGCGGGCGGCT 1499
DB 2223 TTCTCTTTTACTCATGATGATACCAAGTTTGTTCCTTTAAGTGAAGTTAATTCGCGCT 2164
QY 1500 TGGCGTAATCATGATATGCTGTTTCTGTGTGAATTTGTATCCGCTCAATTCAC 1559
DB 2163 TGGCGTAATCATGATATGCTGTTTCTGTGTGAATTTGTATCCGCTCAATTCAC 2104
QY 1560 ACAACATTAAGAGCCGGAAGCATTAAGTAAAGCTGGGGCTTAATGAAGTGAAGCTAAC 1619
DB 2103 ACAACATTAAGAGCCGGAAGCATTAAGTAAAGCTGGGGCTTAATGAAGTGAAGCTAAC 2044
QY 1620 TCACATTAATGCGTTGCGTCACTGCGCCGTTTCCAGTGGGAAACCTGTGTGCCAGC 1679
DB 2043 TCACATTAATGCGTTGCGTCACTGCGCCGTTTCCAGTGGGAAACCTGTGTGCCAGC 1984
QY 1680 TGCATTAATGATGGCCCAACGCGCGGGAGAGAGCGGTTTGGGATTTGGGCGCTTTCCG 1739
DB 1983 TGCATTAATGATGGCCCAACGCGCGGGAGAGAGCGGTTTGGGATTTGGGCGCTTTCCG 1924
QY 1740 CTTCCTGCTCACTGAATCGCTGCGCTCGGTGCTTGGCTGGCGAGCGGATACAGCTC 1799
DB 1923 CTTCCTGCTCACTGAATCGCTGCGCTCGGTGCTTGGCTGGCGAGCGGATACAGCTC 1864
QY 1800 ACTCAAGGCGGTAAACGGTTATCCACAGAAATCAAGGGGATTAACGACAGAAAGACATGT 1859
DB 1863 ACTCAAGGCGGTAAACGGTTATCCACAGAAATCAAGGGGATTAACGACAGAAAGACATGT 1804
QY 1860 GAGCAAAAAGCCGCAAAAAGCCGAGAACCGTAAAGAGCGCGCTGTGTGGCGTTTTC 1919
DB 1803 GAGCAAAAAGCCGCAAAAAGCCGAGAACCGTAAAGAGCGCGCTGTGTGGCGTTTTC 1744
QY 1920 ATAGGCTCGCCGCCCTGTGACGATCACAAAAATCAAGCTGAAGTCAAGGTGGCGAA 1979
DB 1743 ATAGGCTCGCCGCCCTGTGACGATCACAAAAATCAAGCTGAAGTCAAGGTGGCGAA 1684

QY 1980 ACCGCAACGAGACTTAATAAGATACCAAGGCTTTCCCTGGAGAGTCCCTGTGCGCTTC 2039
DB 1683 ACCGCAACGAGACTTAATAAGATACCAAGGCTTTCCCTGGAGAGTCCCTGTGCGCTTC 1624
QY 2040 CTGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTGGGAAAGCGTGG 2099
DB 1623 CTGTTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTGGGAAAGCGTGG 1564
QY 2100 CGCTTTTCATAGCTCAAGCTGATGATCTCAGTTCCGCTGTATAGGTCGTTCCCTCAAGC 2159
DB 1563 CGCTTTTCATAGCTCAAGCTGATGATCTCAGTTCCGCTGTATAGGTCGTTCCCTCAAGC 1504
QY 2160 TGGGCTGTGTGACAGAACCCCGCTTACCGCCGACCCGCTGGGCTTTATCCGTAATCTC 2219
DB 1503 TGGGCTGTGTGACAGAACCCCGCTTACCGCCGACCCGCTGGGCTTTATCCGTAATCTC 1444
QY 2220 GTCTTGAATCCAACCCGCTTAAAGACAGACTTATCCGCACTGGACAGCCGACCTGTAAAC 2279
DB 1443 GTCTTGAATCCAACCCGCTTAAAGACAGACTTATCCGCACTGGACAGCCGACCTGTAAAC 1384
QY 2280 GATTAAGCAGAGAGTATGTAAGCGGTGCTACAGAGTTTGAAGTGTGGCTTAAT 2339
DB 1383 GATTAAGCAGAGAGTATGTAAGCGGTGCTACAGAGTTTGAAGTGTGGCTTAAT 1324
QY 2340 ACGGCTACACTGAAGACAGATTTGTGTATCTGCGCTGTGTAAGCCGTTAACCCTTG 2399
DB 1323 ACGGCTACACTGAAGACAGATTTGTGTATCTGCGCTGTGTAAGCCGTTAACCCTTG 1264
QY 2400 GAAAAAGAGTTGATGCTTGTATCCGGCAAAACAAACCCGCTGTGATGCGGTGTTT 2459
DB 1263 GAAAAAGAGTTGATGCTTGTATCCGGCAAAACAAACCCGCTGTGATGCGGTGTTT 1204
QY 2460 TTGTTTGAACACACATTAACCGGCAAAAAAGATCTCAAGAAAGATCTTTGATCT 2519
DB 1203 TTGTTTGAACACACATTAACCGGCAAAAAAGATCTCAAGAAAGATCTTTGATCT 1144
QY 2520 TTTTCAAGGGGTGACGCTCACTGAGAACGAAACCTCAAGTTAAGGATTTTGTGATGA 2579
DB 1143 TTTTCAAGGGGTGACGCTCACTGAGAACGAAACCTCAAGTTAAGGATTTTGTGATGA 1084
QY 2580 GATTATCAAAAAGATCTTCACTTATGATCTTTTAAATTAATAAGATTTTAATCA 2639
DB 1083 GATTATCAAAAAGATCTTCACTTATGATCTTTTAAATTAATAAGATTTTAATCA 1024
QY 2640 TCTTAAAGTATATAGATACTTGTGATGACAGTTTCAATGCTTATATCAGTAGGAC 2699
DB 1023 TCTTAAAGTATATAGATACTTGTGATGACAGTTTCAATGCTTATATCAGTAGGAC 964
QY 2700 CTATCTCAGGATCTGCTATTTTGTTCATCATAGTTGCTGACTCCCGGTGTGATGA 2759
DB 963 CTATCTCAGGATCTGCTATTTTGTTCATCATAGTTGCTGACTCCCGGTGTGATGA 904
QY 2760 TTAATCAATACGAGAGGCTTACCATCTGGGCCCGAGTCTGCAATGATACCGAGACC 2819
DB 903 TTAATCAATACGAGAGGCTTACCATCTGGGCCCGAGTCTGCAATGATACCGAGACC 844
QY 2820 CAGGCTCACGGGCTCCAGTTTATCAGCAATAAACAGCAGCGGAGAGGCGAGCGCA 2879
DB 843 CAGGCTCACGGGCTCCAGTTTATCAGCAATAAACAGCAGCGGAGAGGCGAGCGCA 784
QY 2880 GAATGTGCTGTGCACTTTATCCGCTCCATCCAGCTTAATTAATTTGGTCCGGGAAAGCTA 2939
DB 783 GAATGTGCTGTGCACTTTATCCGCTCCATCCAGCTTAATTAATTTGGTCCGGGAAAGCTA 724
QY 2940 GAGTAAGTGTGGCCAGTTTAATAGTTTGGCAACGTTGTGCTCAATGCTACAGGCAATCG 2999
DB 723 GAGTAAGTGTGGCCAGTTTAATAGTTTGGCAACGTTGTGCTCAATGCTACAGGCAATCG 664
QY 3000 TGGTGTACGCTGTGTGTTTGTATAGCTTCAATTAAGTCCGGTTCCCAAGATCAAGGC 3059
DB 663 TGGTGTACGCTGTGTGTTTGTATAGCTTCAATTAAGTCCGGTTCCCAAGATCAAGGC 604
QY 3060 GAGTTACATGATCCCCCATGTTTGTGCAAAAAGCGGTTAGCTCTTCCGTCCGATCG 3119

QY	2221	TCTTAGTCCAA	CCCGGTAA	GACAGACTTATCCCACTGGCAGACCACTGTAA	CAG	2280
Db	1442	TCCTAGTCCAA	CCCGGTAA	GACAGACTTATCCCACTGGCAGACCACTGTAA	CAG	2280
QY	2281	GATTACGACGAGGAT	TGTAAGCGGTGCTACAGATTTCTTAAGTGGGCTTA	ACTA	1383	
Db	1382	GATTACGACGAGGAT	TGTAAGCGGTGCTACAGATTTCTTAAGTGGGCTTA	ACTA	1383	
QY	2341	CGGCTACCTAGAAG	ACATATTTGGTATCTCGCTCTGCTGAAACCGCTTA	CTTGG	2400	
Db	1322	CGGCTACCTAGAAG	ACATATTTGGTATCTCGCTCTGCTGAAACCGCTTA	CTTGG	2400	
QY	2401	AAAAAGATTGGT	AGCTTGAATCCGGCAAAACAAACCGCTGGTACCGGTGTTTTT	1203		
Db	1262	AAAAAGATTGGT	AGCTTGAATCCGGCAAAACAAACCGCTGGTACCGGTGTTTTT	1203		
QY	2461	TGTTGCAAGAGAG	ATTACGGCAGAAAAAAGATCTCAAGAAAGATCTTTGATCTT	2520		
Db	1202	TGTTGCAAGAGAG	ATTACGGCAGAAAAAAGATCTCAAGAAAGATCTTTGATCTT	2520		
QY	2521	TTCTACGGGCTCT	GACGCTCAGTGAAGAAACCTCAGTTAAGGATTTTGGTCA	TGAG	2580	
Db	1142	TTCTACGGGCTCT	GACGCTCAGTGAAGAAACCTCAGTTAAGGATTTTGGTCA	TGAG	2580	
QY	2581	ATTATCAAAAAAG	ATTTCACTTGAATCTTTTAAATTAATAAATGAAGTTTAAATCA	TAT	1083	
Db	1082	ATTATCAAAAAAG	ATTTCACTTGAATCTTTTAAATTAATAAATGAAGTTTAAATCA	TAT	1083	
QY	2641	CTTAAGTAAAT	TGATGCTAACTTGGTCTGACAGTTACCAATGCTTAAATGA	AGTTTAAATCA	1023	
Db	1022	CTTAAGTAAAT	TGATGCTAACTTGGTCTGACAGTTACCAATGCTTAAATGA	AGTTTAAATCA	1023	
QY	2701	TATCTACGATCT	GTCTATTTGGTATCATCAATGCTGCTGATCCCGTGTAGAT	2760		
Db	962	TATCTACGATCT	GTCTATTTGGTATCATCAATGCTGCTGATCCCGTGTAGAT	2760		
QY	2761	AACTACGAT	ACGAGGGCTTACACTCTGCCCCAGTGTCAATGATACCGAGACC	2820		
Db	902	AACTACGAT	ACGAGGGCTTACACTCTGCCCCAGTGTCAATGATACCGAGACC	2820		
QY	2821	ACGCTACCGGCT	CTCAGATTTATCAGCAATAAACAGCAGACCGAAGGGCCGAG	CGAG	2880	
Db	842	ACGCTACCGGCT	CTCAGATTTATCAGCAATAAACAGCAGACCGAAGGGCCGAG	CGAG	2880	
QY	2881	AAAGTCTCTG	CACTTTATCCGCTCCATCACTTTAATTTTCCGGGAACTAG	2940		
Db	782	AAAGTCTCTG	CACTTTATCCGCTCCATCACTTTAATTTTCCGGGAACTAG	2940		
QY	2941	AGTAAGT	AGTTGCGCAAGTGTGCAATTTGTTGCGGGAAGCTAG	723		
Db	722	AGTAAGT	AGTTGCGCAAGTGTGCAATTTGTTGCGGGAAGCTAG	723		
QY	3001	GGTGTACGCT	GTGTTGGTATGCTTCAATCACTCGGTGTTCCAAAGATCAAG	CGG	3060	
Db	662	GGTGTACGCT	GTGTTGGTATGCTTCAATCACTCGGTGTTCCAAAGATCAAG	CGG	3060	
QY	3061	AGTTACATGAT	CCCCCATGTTGGCAAAAAAGCGTTAGCTCTTCCGCTCCGATCGT	3120		
Db	602	AGTTACATGAT	CCCCCATGTTGGCAAAAAAGCGTTAGCTCTTCCGCTCCGATCGT	3120		
QY	3121	TGTCAAGAA	TAATGTTGGCGCAGATGTTTACATCACTAGTTAATGCAAGACATGCA	TAATTC	3180	
Db	542	TGTCAAGAA	TAATGTTGGCGCAGATGTTTACATCACTAGTTAATGCAAGACATGCA	TAATTC	3180	
QY	3181	TCCTAAGT	ATGCAATCCGTAAGATGCTTTCTGTGATGCTGAGTACCAACCAAGTC	483		
Db	482	TCCTAAGT	ATGCAATCCGTAAGATGCTTTCTGTGATGCTGAGTACCAACCAAGTC	483		
QY	3241	ATTCTGAGAA	TGTTAATGCGCGACCGAGTGTCTTTGCGCGCGCAATACGGGATTA	3300		
Db	422	ATTCTGAGAA	TGTTAATGCGCGACCGAGTGTCTTTGCGCGCGCAATACGGGATTA	3300		
QY	3301	TACCGGCA	CATGACAACTTTAAAGTCAATCATTTGAAAAAGCTTTCTTGGGGCG	3360		

```

Db      362 TACGCGGCACATGACAGAACTTTAAAGTCTCATTTGAAACCTTTCTCGGGCG 303
Qy      3361 AAAAATCTCAAGAGCTTTACCGCTGTGAGATCCAGTTGATGTAACCACTCGTGCAG 3420
Db      302 AAAATCTCAAGAGCTTTACCGCTGTGAGATCCAGTTGATGTAACCACTCGTGCAG 243
Qy      3421 CAATGATCTTCAGATCTTTTAACTTTTACAGAGCTTTCTGGGTGAGCAAAACAGAGAG 3480
Db      242 CAATGATCTTCAGATCTTTTAACTTTTACAGAGCTTTCTGGGTGAGCAAAACAGAGAG 183
Qy      3481 GCAAAATGCGGCAAAAGAGAAATTAAGGCGACACGAGAAATGTTAAATCACTACTCTT 3540
Db      182 GCAAAATGCGGCAAAAGAGAAATTAAGGCGACACGAGAAATGTTAAATCACTACTCTT 123
Qy      3541 CTTTTTCAATATTAATTAAGACATTTATCAGGGTATTTCTCATGACGGAATCATATTT 3600
Db      122 CTTTTTCAATATTAATTAAGACATTTATCAGGGTATTTCTCATGACGGAATCATATTT 63
Qy      3601 TGATGATTTTGAAGAAATTAACAAATAGGGTTCCGGGCACTTTCCCGAAAGTCC 3660
Db      62 TGAATGATTTTGAAGAAATTAACAAATAGGGTTCCGGGCACTTTCCCGAAAGTCC 3
Qy      3661 AC 3662
Db      2 AC 1

```

```

RESULT 8
US-10-206-030-4/c
; Sequence 4, Application US/10206030
; Publication No. US20020177160A1
; GENERAL INFORMATION:
; APPLICANT: Pagedec, Hal S.
; APPLICANT: Lindebo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/206,030
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: eynthetic construct
US-10-206-030-4

```

```

Query Match      72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```

```

Qy      1 CTAATGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAAGTC 60
Db      3637 CTAATGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAAGTC 3578
Qy      61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTAATAATCAAAAGAAATGACCGA 120
Db      3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTAATAATCAAAAGAAATGACCGA 3518
Qy      121 GATAGGGTGAAGTGTGTTCCAGTTTGAACAAGAGTCCATATTAAGAAGTGAAGTC 180
Db      3517 GATAGGGTGAAGTGTGTTCCAGTTTGAACAAGAGTCCATATTAAGAAGTGAAGTC 3458
Qy      181 CAACGTCAAAAGGCGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACATCAAC 240
Db      3457 CAACGTCAAAAGGCGAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACATCAAC 3398

```

```

Qy      241 CTAATCAAGTTTGTGGGGTGAAGTCCGTAAGCACTAAATCGGAACCTTAAGGAG 300
Db      3397 CTAATCAAGTTTGTGGGGTGAAGTCCGTAAGCACTAAATCGGAACCTTAAGGAG 3338
Qy      301 CCCCCGATTTTGAAGCTTTGACGGGGAAGCCGCGCAAGCTGCGAAGAAAGGAAGGA 360
Db      3337 CCCCCGATTTTGAAGCTTTGACGGGGAAGCCGCGCAAGCTGCGAAGAAAGGAAGGA 3278
Qy      361 AGCGAAAGAGCGGGCGCTAGAGCGCTGGCAAGTGAAGCGGTCAAGCTGCGGTAAAC 420
Db      3277 AGCGAAAGAGCGGGCGCTAGAGCGCTGGCAAGTGAAGCGGTCAAGCTGCGGTAAAC 3218
Qy      421 CACACCGCGCGCTTAATGCGCGCTACAGGGGCGCTCCATTCGCGCATTCAGGCTGCG 480
Db      3217 CACACCGCGCGCTTAATGCGCGCTACAGGGGCGCTCCATTCGCGCATTCAGGCTGCG 3158
Qy      481 CAATGTTGGGAAGGAGGAGTGGTGGGCGCTCTTCCCTATTAAGCCAGCTGGCGAAAG 540
Db      3157 CAATGTTGGGAAGGAGGAGTGGTGGGCGCTCTTCCCTATTAAGCCAGCTGGCGAAAG 3098
Qy      541 GGAATGTGCTGCAAGCGAATTAAGTTGAGTACCCAGGGTTTCCAGTCAACGTTG 600
Db      3097 GGAATGTGCTGCAAGCGAATTAAGTTGAGTACCCAGGGTTTCCAGTCAACGTTG 3038
Qy      601 TAAACGACGCGCGAGTGAAGCGCGCTGTTCAATTCAGTTTGAACCCGAGGAGACGG 660
Db      3037 TAAACGACGCGCGAGTGAAGCGCGCTGTTCAATTCAGTTTGAACCCGAGGAGACGG 2978
Qy      661 GCAGACTGCGGTCGAATGTGTTTACAGCGTATGAGAGCAGATGTCGACAC 720
Db      2977 CCGGGGTGCGG-----CGGCTTAAGAACTAAGTGAATCCCG--CGG 2939
Qy      721 GCTGCAAGAACGCAAGCTAATTAACCTTAAGAAAGATTAATTCATGACGTAA 780
Db      2938 GCTGCAAGAAATCTTAATTTGATTAATTCATTCAGTCAAGTGAATCCCGACAGTTA 2879
Qy      781 AGATTAATCATGCGTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
Db      2878 CAATCTAAGAAAGGACATGTGTCAGCGTTTCGTTGGAGATCTTTCGAAAGGCAAT 2819
Qy      841 TGAGTACCGGCGCCCCCTCGAGGTGCAAGGTATGATTAAGCTTGAATATTCCTGC 900
Db      2818 GTGTGCAAGGTATGATGTGTTGTTGTAAGAAAGAGAGGCGCATCGCAATGAGATATTT 2759
Qy      901 AGCCGCGGGAATCAATGTTCTAGAGCGGCGCACCGCGTGAAGCTCAAGCTTTGT 960
Db      2758 GTTGAATAGGTGCTGTAATGAAGCGATCATCTTCAATGTTGTCGAATTTGAAGT 2699
Qy      961 TCCCTTAAGTGAAGGTTAATTAATCCCATGCGTCAATTTTAAGCAACTATCTTCTAG 1020
Db      2698 TAGCTTGAATTCATTTTGTGTTGTTCTGCGGATGATTAACATTTGTAG--TTATAG 2640
Qy      1021 GATTAATCTAGCTCATCAGATCATATCGTGGGCTTTTTCGCGCTCAGTCATGCGC 1080
Db      2639 TTGTAATCTAGATTTGTGTGCGAAGATTTTTCATCTTCTTAAATCAATACCTTTAAC 2580
Qy      1081 CAAGCTGCGCTATCTGGGCAATCGGAGGAGAAAGCGCGTCTTTTCCGCGAGGTTG 1140
Db      2579 TCGATAGATTAACAAAGGTATCACTTCAAACTTTCAGCAACCGCTTGTAGTTTC 2520
Qy      1141 AAGCGGATGAAGAAAGATTTGCGAGATGACGCTGCTCATTTGAAGCTTGAAGGAAAC 1200
Db      2519 CCGTCATCTTTGAAAGATTAAGTGTCTCTGTATCAATACCTTCGAGGATGAGCATCTTG 2460
Qy      1201 GCAAGTTTACATGATGATTCGGAAGGTGCGCATGACGCTTAAACGTTGAACGTGT 1260
Db      2459 AAAAAGTCAAGCGCTTTCATATGATCCGATTAAGGAAAGCAATTAAGACATTAAGAG 2400
Qy      1261 TCGTTCAAGGCACTCGGGAATACCAATTCGTCGCGCTTTTCCGAGCAACGATTCGGATG 1320
Db      2399 AAGTGAATGACAAAGTTGCGCATGGAACAGTATGTTTCCAGTATGCGAAATTAATTTA 2340
Qy      1321 TCAGCCGGAAGCGCATGAGCAACCGGAACATACCGCGCAACGCGGAACCTGCGTCCG 1380

```

D	2339	AGGGTAAGCTTTCCGATATGTACATCACTTCAACCCCTCCATCGACAGAAAAATTTGTGCG	2280
Q	1381	GTGTGGAGATTATATACAGCGGTCGGCGCTGGGAAATTATACGTACCGAGAGACGGGTATC	1440
D	2279	CCATTACATCAACCATCTAATTTCAACGAATTTGGACACATCTCAGTAAAGATTCTTCT	2220
Q	1441	CTGGCTGGATCCCGCAGAAAAATGGACATACCCCGTAGTTAACCCGGCGGGCGGCTT	1500
D	2219	C---CTTTACTCATCGGTACCAGCTTTTGTTCCTTTAGTAGAGGGTTAATTGGCGGCTT	2163
Q	1501	GGCGTAATCAATGACATAGCTGTTTCTGTGTGAATTTGTATCCGGTCCAAATTCACAA	1560
D	2162	GGCGTATCATGTCTACATACGTGTTCCTGTGTGAATTTGTATCCGGTCCAAATTCACAA	2103
Q	1561	CAACATTCAGACCCGGAGAGCATTAATGTAAAGCCCTGGGGTGTCCCTAATAGTAGCTAAT	1620
D	2102	CACATATACAGCCCGGAGAGCATTAATGTAAAGCCCTGGGGTGTCCCTAATAGTAGCTAAT	2043
Q	1621	CACATTAATTTGGCTGTGCTCACTGACCCCTTTTCAGTCGGGAAACCTGTGGTCAAGCT	1680
D	2042	CACATTAATTTGGCTGTGCTCACTGACCCCTTTTCAGTCGGGAAACCTGTGGTCAAGCT	1983
Q	1681	GCATTAAATGAATCGGCCCAACGCGCGGGAGAGCGGTTTGGCTAATTTGGCGCTTTCCGC	1740
D	1982	GCATTAAATGAATCGGCCCAACGCGCGGGAGAGCGGTTTGGCTAATTTGGCGCTTTCCGC	1923
Q	1741	TTTCCTCGTCACTGACTCGCTGCGCTGGTTCGCTTGGCTGGCGGAGCGGTATCAGCTCA	1800
D	1922	TTTCCTCGTCACTGACTCGCTGCGCTGGTTCGCTTGGCTGGCGGAGCGGTATCAGCTCA	1863
Q	1801	CTCAAAAGCGGTAAATACGTTTATCCACAGAAATCAGGGGATPACGAGGAAAGAAACATGTG	1860
D	1862	CTCAAAAGCGGTAAATACGTTTATCCACAGAAATCAGGGGATPACGAGGAAAGAAACATGTG	1803
Q	1861	AGCAAAAGGCCACGAAAGGCCAGGAACCGTAAAAAGCCGCTTGTCTGGCGTTTTCACA	1920
D	1802	AGCAAAAGGCCACGAAAGGCCAGGAACCGTAAAAAGCCGCTTGTCTGGCGTTTTCACA	1743
Q	1921	TAGGCTCCGCCCCCTCAGCAGCATCAAAAATGCAGCCTCAAGTCAAGAGTGGCGAAA	1980
D	1742	TAGGCTCCGCCCCCTCAGCAGCATCAAAAATGCAGCCTCAAGTCAAGAGTGGCGAAA	1683
Q	1981	CCCGCAGAGCATATAAAGATACAGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTTCC	2040
D	1682	CCCGCAGAGCATATAAAGATACAGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTTCC	1623
Q	2041	TGTTCCCAACCTTCGCGCTTACCGGATACCTGTGCGCGCTTTCCTCTCGGAAAGCGTTCG	2100
D	1622	TGTTCCCAACCTTCGCGCTTACCGGATACCTGTGCGCGCTTTCCTCTCGGAAAGCGTTCG	1563
Q	2101	GCTTTCTCAATAGTCAACGCTGTAAGTATCTCAGTTCCGTTAGGTGTGTTCCCTCAAGT	2160
D	1562	GCTTTCTCAATAGTCAACGCTGTAAGTATCTCAGTTCCGTTAGGTGTGTTCCCTCAAGT	1503
Q	2161	GGGCTGTGTGGACCAACCCCGTTTCAGCCCGACCGGTGCGCTTATCCGTAATCTATCG	2220
D	1502	GGGCTGTGTGTGGACCAACCCCGTTTCAGCCCGACCGGTGCGCTTATCCGTAATCTATCG	1443
Q	2221	TTTTAAGTCCAAACCCGGTAAAGACACACATTAATGCGCATGCGACAGCAGCTGTATACAG	2280
D	1442	TTTTAAGTCCAAACCCGGTAAAGACACACATTAATGCGCATGCGACAGCAGCTGTATACAG	1383
Q	2281	GATTAGCAGACGAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTGGCCCTTAATA	2340
D	1382	GATTAGCAGACGAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTGGCCCTTAATA	1323
Q	2341	CGGCTACACTAAGAGACAGTATTTTGTATCGGCTCTGTCTGAAGCAGTTACCTTGG	2400
D	1322	CGGCTACACTAAGAGACAGTATTTTGTATCGGCTCTGTCTGAAGCAGTTACCTTGG	1263
Q	2401	AAAAAGATTGTAGCTTTGATCCGGCAACCAACCGCTGGTAGAGGGTGTCTTTT	2460

Db 1262 AAAAAAGATGGTAGCTCTTATCCGGCAAAACAACCGCGTGGTAGCGGCTTTT 1203

QY 2461 TGTGTGCAAGCAGCAAGATTACGCGACAAAAAAGATCTCAAGAAATCTTTGATCTT 2520

Db 1202 TGTTTGCAAGCAGCAAGATTACGCGAGAAAAAAGATCTCAAGAAATCTTTGATCTT 2520

QY 2521 TTCTTGGGGGTCTGACCTCAGTGAAGAAACAACTCACTTAAGGATTTTGGTCAATG 1143

Db 1142 TCTACGGGGGTCTGACCTCAGTGAAGAAACAACTCACTTAAGGATTTTGGTCAATG 1083

QY 2581 ATTTACAAAAGATCTTACCTAATCTTTTAAATTAAAAATGAAGTTTAAATCAAT 2640

Db 1082 ATTTACAAAAGATCTTACCTAATCTTTTAAATTAAAAATGAAGTTTAAATCAAT 1023

QY 2641 CTAAAGATTATAGATTAACTTGGTCTGACAGTTACCAATGCTTAAATCACTGAGCAAC 2700

Db 1022 CTAAAGATTATAGATTAACTTGGTCTGACAGTTACCAATGCTTAAATCACTGAGCAAC 2700

QY 2701 TATCTCAGGATCTGTCTAATTCGTTCATCAATAGTTGCTTGAATCCCGCTGTGATG 2760

Db 962 TATCTCAGGATCTGTCTAATTCGTTCATCAATAGTTGCTTGAATCCCGCTGTGATG 2760

QY 2761 AACTACGATACGGAGGGCTTACATCTGCCCCCAGTGTGCAATACCGCAACC 2820

Db 902 AACTACGATACGGAGGGCTTACATCTGCCCCCAGTGTGCAATACCGCAACC 2820

QY 2821 ACGCTACCGGCTTCAGATTATACGACATAAACAGCAGACCGGAAGGCTCAGCGAG 2880

Db 842 ACGCTACCGGCTTCAGATTATACGACATAAACAGCAGACCGGAAGGCTCAGCGAG 2880

QY 2881 AAGTGTCTGCAACTTTATCCGCTCCATCCATCTAATTAATTTGTGCGGGAAAGCTAG 2940

Db 782 AAGTGTCTGCAACTTTATCCGCTCCATCCATCTAATTAATTTGTGCGGGAAAGCTAG 2940

QY 2941 AGTAAGTATGTCAGCACTTAATCTTTGCGCAAGTTGTGCCATTTGTCAGCAGGATGT 723

Db 722 AGTAAGTATGTCAGCACTTAATCTTTGCGCAAGTTGTGCCATTTGTCAGCAGGATGT 3000

QY 3001 GGTGTACGGCTGTCTTTGGTATGAGCTTCAATCAGCTCCGGTTCCCAAGATCAAGCG 663

Db 662 GGTGTACGGCTGTCTTTGGTATGAGCTTCAATCAGCTCCGGTTCCCAAGATCAAGCG 3060

QY 3061 AGTTACATGATCCCAATGTTGTGCAAAAAAAGCGTTAGCTTCCGATCCCGATCGT 603

Db 602 AGTTACATGATCCCAATGTTGTGCAAAAAAAGCGTTAGCTTCCGATCCCGATCGT 3120

QY 3121 TGTGAGAAGTAAGTGGCGCGAGTGTATCACTGATGGTTATGGCAGACTGATTAATTC 543

Db 542 TGTGAGAAGTAAGTGGCGCGAGTGTATCACTGATGGTTATGGCAGACTGATTAATTC 483

QY 3181 TCTTACGTGATCCCATCCGTAAGATGCTTTCTGTGATCGTGAATCTCAACCAAGTC 3240

Db 482 TCTTACGTGATCCCATCCGTAAGATGCTTTCTGTGATCGTGAATCTCAACCAAGTC 3240

QY 3241 ATTCTGAAATAGTGTATGCGGACAGCAATCTTCTCCCGGCTCAATACCGGATTA 3300

Db 422 ATTCTGAAATAGTGTATGCGGACAGCAATCTTCTCCCGGCTCAATACCGGATTA 3300

QY 3301 TACCGCGCATATGCAAACTTTAAAGTGTCAATGTGAAAAAGTCTTCTGGGGCG 363

Db 362 TACCGCGCATATGCAAACTTTAAAGTGTCAATGTGAAAAAGTCTTCTGGGGCG 3660

QY 3361 AAAATCTCAAGATCTTACCGCTGTGAGATCCAGTTCCATGTACCACTGTGACCC 3420

Db 302 AAAATCTCAAGATCTTACCGCTGTGAGATCCAGTTCCATGTACCACTGTGACCC 3420

QY 3421 CAATGATCTTCAAGATCTTCTTACTTCAACAGGTTTCTGGGTGACAAAAACAGGAG 3480

Db 242 CAATGATCTTCAAGATCTTCTTACTTCAACAGGTTTCTGGGTGACAAAAACAGGAG 3480

QY 3481 GCAAAAATGCGCAAAAAAGCAATTAAGGGGACACAGGAATGTGAAATATCTCTT 183

Db 182 GCAAAAATGCGCAAAAAAGCAATTAAGGGGACACAGGAATGTGAAATATCTCTT 3540

182 GCAAAAATGCGCAAAAAAGCAATTAAGGGGACACAGGAATGTGAAATATCTCTT 123

Query 3541 CCTTTTGAATATTTAGACATTTATGAGGTTATTTGTCATGACGGAATACATAT 3600
Db 122 CCTTTTGAATATTTAGACATTTATGAGGTTATTTGTCATGAGGATACATAT 63
Query 3601 TGAATGATTTAGAAAATAAACAATAGGGTCCGCGACATTTCCCGAAAGTGCC 3660
Db 62 TGAATGATTTAGAAAATAAACAATAGGGTCCGCGACATTTCCCGAAAGTGCC 3
Query 3661 AC 3662
Db 2 AC 1
RESULT 9
US-10-211-079-4/c
Sequence 4, Application US/10211079
Publication No. US20030148315A1
GENERAL INFORMATION:
APPLICANT: Padgett, Hal S.
APPLICANT: Vaewhongs, Andrew A.
APPLICANT: Vojdani, Fakhrieh S.
APPLICANT: Smith, Mark L.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
FILE REFERENCE: P-IG 5381
CURRENT APPLICATION NUMBER: US/10/211,079
PRIORITY FILING DATE: 2002-08-01
PRIORITY APPLICATION NUMBER: US 10/098,155
PRIORITY FILING DATE: 2002-03-14
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 3637
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
US-10-211-079-4
Query Match 72.8%; Score 2664.8; DB 15; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
Query 1 CTTAAATGTAAGCGTTAATATTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60
Db 3637 CTTAAATGTAAGCGTTAATATTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578
Query 61 ATTTTAAACAATAGCGCGAAATCGGCAAAATCCCTATAATCAAAAGAAATGACCGA 120
Db 3578 ATTTTAAACAATAGCGCGAAATCGGCAAAATCCCTATAATCAAAAGAAATGACCGA 3518
Query 121 GATAGGGTGAAGTGTGTTCCAGTTTGGACAAGATCCACTATTTAAAGACGTGACATC 180
Db 3517 GATAGGGTGAAGTGTGTTCCAGTTTGGACAAGATCCACTATTTAAAGACGTGACATC 3458
Query 181 CAAGCTCAAGGCGCAAAACCGCTTATCAGGCGGATGCCCACTAGCTGAACCATCAC 240
Db 3457 CAAGCTCAAGGCGCAAAACCGCTTATCAGGCGGATGCCCACTAGCTGAACCATCAC 3398
Query 241 CTAATCAAGTTTGTGGGTGAGAGTGCGCTTAAGCACTAAATCGGAACCTTAAAGGAG 300
Db 3397 CTAATCAAGTTTGTGGGTGAGAGTGCGCTTAAGCACTAAATCGGAACCTTAAAGGAG 3338
Query 301 CCCCCGATTTAGAGCTTGACGCGGCAAAACCGCGCAACCTGCGAGAAAGGAGAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGCGGCAAAACCGCGCAACCTGCGAGAAAGGAGAGAA 3278
Query 361 AGCGAAAGAGAGCGGCGCTTAGGCGCTGCGCAAGTGTACGGGTACGCGTCCGCTAAC 420
Db 3277 AGCGAAAGAGAGCGGCGCTTAGGCGCTGCGCAAGTGTACGGGTACGCGTCCGCTAAC 3218
Query 421 CACACCCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCCATTCGCAATTCAGGCTGCG 480

Db 3217 CACACCCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCCATTCGCAATTCAGGCTGCG 3158
Query 481 CAACTGTTGGAGAGCGCATGCGTGGCGGCTCTTCGCTATTTAGCCAGCTGGCGAAG 540
Db 3157 CAACTGTTGGAGAGCGCATGCGTGGCGGCTCTTCGCTATTTAGCCAGCTGGCGAAG 3098
Query 541 GGGATGTCGTCAGAGCGATTAAGTGGGTAAAGCCAGGGTTTCCAGTACAGACGTTG 600
Db 3097 GGGATGTCGTCAGAGCGATTAAGTGGGTAAAGCCAGGGTTTCCAGTACAGACGTTG 3038
Query 601 TAAACGACGCGCAGTGAAGCGCGCTGCTTATTCAGCTTTTGAACCCGAGAGAGAG 660
Db 3037 TAAACGACGCGCAGTGAAGCGCGCTGCTTATTCAGCTTTTGAAGCCGAGAGAG 2978
Query 661 GCAGACTCGCGTGCAGAAATGTGTTTAAAGCGATGAGAGAGATGCTGCACAC 720
Db 2977 CCGCGGTGCGG-----CGCTCTAGAACTAGTGAATCCCC--CGG 2939
Query 721 GCTGCAGAAACGACGATGATTAACCTAGAAAGATATATATTTGTGACGTACGTTA 780
Db 2938 GCTGCAGAAATCTTATTTGTATGTATTCATCCATGCGCATGTGTATCCAGACAGATTA 2879
Query 781 AGATATCATGCTTAAATTTAGACGATGCGATCTGTATATACGCTCATATAGGCGCAT 840
Db 2878 CAACTCAAGAGAGACATGTGTCAGCTTTTGTGGATCTTTTGAAGGCGCATAT 2819
Query 841 TGCGTACCGGCGCCCCCTCGAGGTGACGCGTATCGATTAAGCTTATTCGATTCCTGC 900
Db 2818 GTGTCAGAGGTATGTTGTCTGTGTAAGAGACAGGCGCATTCGCAATGTGATATTTT 2759
Query 901 AGCCCGGGGATTCATGATTTCTAGAGCGCGCGCACCGCGGTGAGTCCAGCTTTGT 960
Db 2758 GTTATATATGCTGCTAGTGAACGATCATCTTCAATGTTGTGCGAATTTTGAAGT 2699
Query 961 TCCCTTAGTAGAGGTTAATATGATCCCATGCGCAATTTTACGACGATCTTTCTT 1020
Db 2698 TAGCTTTGATTCATTTCTTTGTTGTCGCGGATGTATCACTGTGTAG--TTATAG 2640
Query 1021 GGTATATCTAGCTGATCAAGATCATATCGTGGGCTTTTTCGCGCTCATGATCGCC 1080
Db 2639 TTGATCTGAGTTGTGTGCGAGATGTTTCATCTTTTAAATCAATACCTTTTAAAC 2580
Query 1081 CAACTGCGCTATCTGCGCATCGGAGAGAGAAAGCCGCGCTTTTCCGCAAGTTG 1140
Db 2579 TCGATTAAGATTAACAGGATATCACTTCAAACTTGCACAGACCGCTTTGTAGTTC 2520
Query 1141 AAGCGCATGGAAGAGTTGCGAGATGACCTGCTGATGACGTTGAGCGAATAAC 1200
Db 2519 CCGTATCTTTGAAGATATATGTCGTTCTGTATCATTAACCTTGGCGATGACCTTTG 2460
Query 1201 GCACGTTTACCATGATATTCGGAAGGTGCGCATGACGCGCTTTAACGSGTAACGT 1260
Db 2459 AAAAAGTACGCGCTTATATATGATTCGGAATACGGAAAGCATTAACCATTAAGAG 2400
Query 1261 TCGTTACGCGCACTGGAATACAGTTGTCGCGGCTTTTCCGACACAGTTCCGATGG 1320
Db 2399 AAGGTATGACAGATGTTGCGCATGGAACAGTATTTTCCAGTATGCAAAATTAATTA 2340
Query 1321 TCAGCCGAGCGCATCAGCAACCGCAACAATACCGCGACAGCGGAACTGCGTCCG 1380
Db 2339 AGGTTAGCTTTTCCGATGATGATCACTTACCTCTCCATGACAGAAATTTGTC 2280
Query 1381 GTGTGAGATTAATGACAGCGGTGCGCGCTGAGATTAAGTCAAGCAGAGAGCGGATATC 1440
Db 2279 CCATTAACATCAGCATTAATTAACAAGAAATGGGACAACTCAGTGAAGAAAGTTCTTCT 2220
Query 1441 CTGCTGATGCGCAGAAATGACATGATATCCCGTGAGTTACCGCGCGGCGGCTT 1500
Db 2219 C---CTTACTCATCGGTACCCAGCTTTTGTCCCTTTAGTGAAGGTTAAATGCGCTT 2163
Query 1501 GCGCTAATCATAGGCTAGCTGTTCTGTGTGAATTTGTTATCGGCTCAAAATTCACA 1560


```
Db 2162 GGGCTAATCATGCTCATAGCTTTCTGTGTGAATTTGTAATCCGCTCACAATTCACACA 2103
Qy 1561 CAACATTCGAGCGCGAAGCATAAAGTAAAGCCCTGGGGTCCCTTAATGAGTGAAGCTAACT 1620
Db 2102 CAACATTCGAGCGCGAAGCATAAAGTAAAGCCCTGGGGTCCCTTAATGAGTGAAGCTAACT 1620
Qy 1621 CACATTAATTTGGGTTGGGCTCACTGCGCTTTTCAGTTCGGGAAACCTGTCCGTCCAGCT 2043
Db 2042 CACATTAATTTGGGTTGGGCTCACTGCGCTTTTCAGTTCGGGAAACCTGTCCGTCCAGCT 1983
Qy 1681 GGATTAATGAATCGGCAACCGCGGGGAGAGGGGCTTTCGTAATTTGGCGCTTTCCGC 1740
Db 1982 GGATTAATGAATCGGCAACCGCGGGGAGAGGGGCTTTCGTAATTTGGCGCTTTCCGC 1923
Qy 1741 TTCTCGCTCACTGACTCGCTGCGCTGCTGCTTTCGCTGCGCGAGCGAGTTCAGCTCA 1800
Db 1922 TTCTCGCTCACTGACTCGCTGCGCTGCTGCTTTCGCTGCGCGAGCGAGTTCAGCTCA 1863
Qy 1801 CTCAAGGCGGTTAATACGTTTATCCAGAAATCCAGAGGATTAACGAGAAAGAAATCATGTG 1860
Db 1862 CTCAAGGCGGTTAATACGTTTATCCAGAAATCCAGAGGATTAACGAGAAAGAAATCATGTG 1803
Qy 1861 AGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGCGCTGCTGCGGCTTTTCCA 1920
Db 1802 AGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGCGCTGCTGCGGCTTTTCCA 1743
Qy 1921 TAGGCTCGGCGCGCGCTGAGAGCATCACAAAATCCAGCTCAAGTCAGAGTGGCGAAA 1980
Db 1742 TAGGCTCGGCGCGCGCTGAGAGCATCACAAAATCCAGCTCAAGTCAGAGTGGCGAAA 1981
Qy 1981 CCCGACAGACTATTAAGATACAGAGGCTTTCCCGTGAAGCTCCCTGTCGCTCC 2040
Db 1682 CCCGACAGACTATTAAGATACAGAGGCTTTCCCGTGAAGCTCCCTGTCGCTCC 1623
Qy 2041 TGTTCGACCCCTGCGCTTACCGGATCTGTGCTCCCTTCTCCCTGGGAAAGGTGGC 2100
Db 1622 TGTTCGACCCCTGCGCTTACCGGATCTGTGCTCCCTTCTCCCTGGGAAAGGTGGC 1563
Qy 2101 GCTTTCATATGCTCACGCTGATGATCTCAGTTCCGAGTGGTGGTGGCTTCAAGCT 2160
Db 1562 GCTTTCATATGCTCACGCTGATGATCTCAGTTCCGAGTGGTGGTGGCTTCAAGCT 1503
Qy 2161 GGGCTGTGTCAAGAACCCCGGTTACGCGGACCGCTGCGCTTATCCGTAATCTAATG 2220
Db 1502 GGGCTGTGTCAAGAACCCCGGTTACGCGGACCGCTGCGCTTATCCGTAATCTAATG 1443
Qy 2221 TCTTGAATCAACCCGCTTAAGACAGACTTATCCCACTGGGCGAGCCACTGTGAACAG 2280
Db 1442 TCTTGAATCAACCCGCTTAAGACAGACTTATCCCACTGGGCGAGCCACTGTGAACAG 2281
Qy 2281 GATTAGCAGAGCAGATGTAGGCGGTCTACAGATTCTTGAAGTGGTGGCTTAACGA 2340
Db 1382 GATTAGCAGAGCAGATGTAGGCGGTCTACAGATTCTTGAAGTGGTGGCTTAACGA 1323
Qy 2341 CGGCTACACTGAAGGACATATTTGGTATCTGCGCTGCTGGAAGCCGTTACTTGGG 2400
Db 1322 CGGCTACACTGAAGGACATATTTGGTATCTGCGCTGCTGGAAGCCGTTACTTGGG 2401
Qy 2401 AAAAAGGTTGTAAGCTTTGATCCGCAAAACAAACCGCTGGTAGCGGTGTTTTT 2460
Db 1262 AAAAAGGTTGTAAGCTTTGATCCGCAAAACAAACCGCTGGTAGCGGTGTTTTT 1203
Qy 2461 TGTTCGACGACAGATTAACGCGCAAAAGAAAGATCTCAAGAAATCTTGTGATCTT 2520
Db 1202 TGTTCGACGACAGATTAACGCGCAAAAGAAAGATCTCAAGAAATCTTGTGATCTT 1143
Qy 2521 TTCTACGCGGCTGACGCTCAGTGAAGCAAAACCTCAGTAAAGGATTTTGTGATGAG 2580
Db 1142 TTCTACGCGGCTGACGCTCAGTGAAGCAAAACCTCAGTAAAGGATTTTGTGATGAG 2581
Qy 2581 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATTAAGATTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATTAAGATTTTAAATCAAT 1023

Qy 2641 CTAAAGTATTAATGAGTAAACCTTGCTGACAGTTAACAAATGCTTAATGATGAGCAC 2700
Db 1022 CTAAAGTATTAATGAGTAAACCTTGCTGACAGTTAACAAATGCTTAATGATGAGCAC 963
Qy 2701 TATCTGAGGATCTGTCTAATTTGCTTACATCAATAGTTGCTGACCTCCGCTGTAGAT 2760
Db 962 TATCTGAGGATCTGTCTAATTTGCTTACATCAATAGTTGCTGACCTCCGCTGTAGAT 903
Qy 2761 AACTAGATACGAGAGGCTTACATCTGAGCCCACTGCTGCAATGATACCGAGAACCC 2820
Db 902 AACTAGATACGAGAGGCTTACATCTGAGCCCACTGCTGCAATGATACCGAGAACCC 843
Qy 2821 ACGCTACCGGCTTCAATTTATCAGCAATTAACAGCAACCGGAGAGGCGGAGCCAG 2880
Db 842 ACGCTACCGGCTTCAATTTATCAGCAATTAACAGCAACCGGAGAGGCGGAGCCAG 783
Qy 2881 AAGTGTCTGCAACTTATCCGCTCCATCCAGTCTAATTAATTTGGCGGAGGTAG 2940
Db 782 AAGTGTCTGCAACTTATCCGCTCCATCCAGTCTAATTAATTTGGCGGAGGTAG 723
Qy 2941 AGTAAGTATTTGCCAGTAAATAGTTTGGCAAGCTTTCAGAGCATGCT 3000
Db 722 AGTAAGTATTTGCCAGTAAATAGTTTGGCGAAGCTTTCAGAGCATGCT 663
Qy 3001 GGTGTACGCTGTGCTTGTATGCTTCAATTCAGCTCCGATTCCCAAGATCAAGCG 3060
Db 662 GGTGTACGCTGTGCTTGTATGCTTCAATTCAGCTCCGATTCCCAAGATCAAGCG 603
Qy 3061 AGTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGCTCTTCGCTCCGATGCT 3120
Db 602 AGTACATGATCCCGCATGTTGTGCAAAAAGCGGTAGCTCTTCGCTCCGATGCT 543
Qy 3121 TGTCAAGATAGTTGCGCAGTGTATCACTCATGTTATGAGCAGACATCAATTTTC 3180
Db 542 TGTCAAGATAGTTGCGCAGTGTATCACTCATGTTATGAGCAGACATCAATTTTC 483
Qy 3181 TCTTACGTATGATCCGTAAGATGCTTTTGTGTGACTGTGATGATGATCAACAGATC 423
Db 482 TCTTACGTATGATCCGTAAGATGCTTTTGTGTGACTGTGATGATGATCAACAGATC 3240
Qy 3241 ATTCTGAAATAGTGTATGCGGACAGAGTGTGCTTTGCGCGGCTCAATACGAGATTA 3300
Db 422 ATTCTGAAATAGTGTATGCGGACAGAGTGTGCTTTGCGCGGCTCAATACGAGATTA 363
Qy 3301 TACCGGACCATATACAGAACTTTAAAGTCTATCTATTTGAAAAAGTTCTTGGGGG 3360
Db 362 TACCGGACCATATACAGAACTTTAAAGTCTATCTATTTGAAAAAGTTCTTGGGGG 303
Qy 3361 AAAAATCTCAAGATCTTACCGCTGTGAGATCCAGTGTGATTAACCACTCGTGACC 3420
Db 302 AAAAATCTCAAGATCTTACCGCTGTGAGATCCAGTGTGATTAACCACTCGTGACC 243
Qy 3421 CAATGATCTTACGATCTTTTACTTTCACAGCTTTCTGTGTGAGCAAAACAGGAG 3480
Db 242 CAATGATCTTACGATCTTTTACTTTCACAGCTTTCTGTGTGAGCAAAACAGGAG 183
Qy 3481 GCAAAATGCGGCAAAAGGAAATAGGCGACAGCAAGATTTGATATCTATCTT 3540
Db 182 GCAAAATGCGGCAAAAGGAAATAGGCGACAGCAAGATTTGATATCTATCTT 123
Qy 3541 CTTTTTCAATATTAATGAAGATTTATCAGGTTATTTGTCTATGAGCGGATATAT 3600
Db 122 CTTTTTCAATATTAATGAAGATTTATCAGGTTATTTGTCTATGAGCGGATATAT 63
Qy 3601 TGAATGATTTTGAAGAAATTAACAAATAGGAGTTCCGCGACATTTCCCGAAAAAGTGC 3660
Db 62 TGAATGATTTTGAAGAAATTAACAAATAGGAGTTCCGCGACATTTCCCGAAAAAGTGC 3661
Qy 3661 AC 3662
Db 2 AC 1
```

RESULT 10
US-10-356-708-17/C
Sequence 17, Application US/10356708
Publication No. US20030157682A1
GENERAL INFORMATION:
APPLICANT: Large Scale Biology Corporation
TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
FILE REFERENCE: P-IG 10100
CURRENT APPLICATION NUMBER: US/10/356,708
PRIOR FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: 60/353,722
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 10/098,155
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 10/211,079
PRIOR FILING DATE: 2002-08-01
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 17
LENGTH: 3637
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Encodes cycle 3 GFP
US-10-356-708-17.

Query Match 72.8%; Score 2664.8; DB 15; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY 1 CTAATTGTAAGCCGTAATATTTGTTAAATGCGCTAAATTTTGTAAATCACTG 60
DB 3637 CTAATGTAAGCCGTAATATTTGTTAAATGCGCTAAATTTTGTAAATCACTG 3578
QY 61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTAATTAATCAAAAGATAGACGA 120
DB 3577 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTAATTAATCAAAAGATAGACGA 3518
QY 121 GATAGGGTGAAGTGTGTTCCAGTTTGGACAAAGATCCACTATTAAGAAAGTGACTC 180
DB 3517 GATAGGGTGAAGTGTGTTCCAGTTTGGACAAAGATCCACTATTAAGAAAGTGACTC 3458
QY 181 CAAGTCAAAAGGGGAAACCCGTTATCAGAGGCGATGCGCCCACTAGCTGAACCATCACC 240
DB 3457 CAAGTCAAAAGGGGAAACCCGTTATCAGAGGCGATGCGCCCACTAGCTGAACCATCACC 3398
QY 241 CTAATCAAGTTTGTGGGCTGAGGTCGCTAAAGCACTAAATCGGAACCTTAAGGAG 300
DB 3397 CTAATCAAGTTTGTGGGCTGAGGTCGCTAAAGCACTAAATCGGAACCTTAAGGAG 3338
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAACCCGCGCAACGTCGCAAGAAAGAAAGGAA 360
DB 3337 CCCCCGATTTAGAGCTTGAACGGGGAACCCGCGCAACGTCGCAAGAAAGAAAGGAA 3278
QY 361 AGCGAAAGAGAGCGGCGCTAGGCGCTGCGCAAGTGTACGGCTCAAGCTGCGCGTAAACAC 420
DB 3277 AGCGAAAGAGAGCGGCGCTAGGCGCTGCGCAAGTGTACGGCTCAAGCTGCGCGTAAACAC 3218
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCCATTCGCAATTCAGAGCTGCG 480
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGCTCCCATTCGCAATTCAGAGCTGCG 3158
QY 481 CAACGTGTGGAGAGGCGATGCTGCGGCGCTTTCGCTAATTAAGCGCAAGCTGCGGAAG 540
DB 3157 CAACGTGTGGAGAGGCGATGCTGCGGCGCTTTCGCTAATTAAGCGCAAGCTGCGGAAG 3098
QY 541 GCGATGTCGCGCAAGCGATTAAGTGTGTAACGCAAGGCTTTTCCAGTCAAGAGCTTG 600
DB 3097 GCGATGTCGCGCAAGCGATTAAGTGTGTAACGCAAGGCTTTTCCAGTCAAGAGCTTG 3038
QY 601 TAAACGACGCGCAGTGAAGCGCGCTCGTTCAATTCAGCTTTTGAACCGTGAAGAGCG 660

DB 3037 TAAACGACGCGCAGTGAAGCGCGCTCGTTCAATTCAGCTATTAAGGCGCAATTTGAGCTCCA 2978
QY 661 GCAAGCTCGCGTCAAAATGTTTATACAGCTGATGAGCAAGTGAAGATGCTGACAC 720
DB 2977 CCGCGGTGCGG-----CCGCTCAGAACTAGTGAATGAGTCACTCCG 2939
QY 721 GCTCAGAAACGAGAGTGAATTAACCTGAAGAAATTAATTAATTTGAGAGTACGTTAA 780
DB 2938 GCTCAGAAATTTCTTAATTTGATAGTTATTCATCAAGTGAATGATCCAGAGAGTTA 2879
QY 781 AGATATCATGCTGTAATTAATGACCATGCGATCTGTAAATACAGCTCATATAGGCGCAAT 840
DB 2878 CAATCTAAGAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2819
QY 841 TGGGTACCGGCG 900
DB 2818 GTGTGACAGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2759
QY 901 AGCCGCGGCGATTCAGTATTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
DB 2758 GTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2699
QY 961 TCCCTTATGAGGAGTAAATTAATGATCCATGCGTCAATTTTACAGAGTATCTTCTAG 1020
DB 2698 TAGCTTTGATTCATTCCTTTTGTGTCGCGGATGATGATGATGATGATGATGATGATGAT 2640
QY 1021 GGTAAATCTAGCTGATCAGATCATATTCGTGCGGCTTTTTCGCGCTCATGCTGCGC 1080
DB 2639 TTGTACTGAGTGTGTGTCGAGAAATGTTTCCATCTTTTAAATCAATATCTTTTAAAC 2580
QY 1081 CAAGCTGCGCTATCTGCGGATTCGCGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1140
DB 2579 TCGATACGATTAACAAAGGATTCACCTTCAAACTTGAACGATGATGATGATGATGAT 2520
QY 1141 AAGCGGATGAAAGATTTGCGCGAGATGATCTGCTGCTGATGATGATGATGATGATGAT 1200
DB 2519 CCGTATCTTTGAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
QY 1201 GCACGTTTACATGATATTCGGAAGTGTGATGATGATGATGATGATGATGATGATGAT 1260
DB 2459 AAAAGTCAATGCTTTCATATGATCCGGAATACGGAAGAAAGCAATTAACCATTAAG 2400
QY 1261 TCGTCAAGGCACTCGGATATCAAGTTCGTCGCGCTTTCCGACACAGTTCGGAATG 1320
DB 2399 AAAGTATGACAAAGTGTGCGCATGAAACAGATATTTTCCAGATGTCAAATTAATTA 2340
QY 1321 TCAGCGGAGAGCGATCAGCAACCGGAACAATACCGGCAAGCGGGAATGCGCGCG 1380
DB 2339 AGGATAGCTTTCCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
QY 1381 GTGTGAGATTAATGACAGCGGTGCGGCTGAGATTAATGTCAGCGAGAGACGGATATC 1440
DB 2279 CCAATTAATCAATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220
QY 1441 CTGCGTGAATGCGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 2219 C---CTTATCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2163
QY 1501 GCGCTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 2162 GCGCTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103
QY 1561 CAACATGACAGCGGAAAGATTAAGTGAAGCTGCGGATGATGATGATGATGATGATGAT 1620
DB 2102 CAACATGACAGCGGAAAGATTAAGTGAAGCTGCGGATGATGATGATGATGATGATGAT 2043
QY 1621 CACATTAATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 2042 CACATTAATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1983
QY 1681 GCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1982 GCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923

CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: US/10/066,390
PRIOR FILING DATE: 2002-02-01
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 3637
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
US-10-205-772-4

Query March 72.8%; Score 2664.8; DB 15; Length 3637;

Best Local Similarity 83.8%; Pred. No. 0; Mismatches 567; Indels 25; Gaps 4;

Matches 3070; Conservative 0; Pseudomatches 567; Indels 25; Gaps 4;

QY 1 CTAATTGTAGAGCTTAATTGTTGTTAAATTCGCTTAATTTTGTAAATCAGCTC 60
DB 3637 CTAATTGTAGAGCTTAATTGTTGTTAAATTCGCTTAATTTTGTAAATCAGCTC 3578
QY 61 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTAATAATCAAAAGATAGACGA 120
DB 3577 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTAATAATCAAAAGATAGACGA 3518
QY 121 GATAGGGTGTAGTGTGTTCCAGTTTGAACAAGTCACTATTAAAGACGTGACTC 180
DB 3517 GATAGGGTGTAGTGTGTTCCAGTTTGAACAAGTCACTATTAAAGACGTGACTC 3458
QY 181 CAAGCTCAAAAGGCGAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACC 240
DB 3457 CAAGCTCAAAAGGCGAAACCGTCTATCAGGGCGATGGCCCACTAGCTGAACCATCACC 3358
QY 241 CTAATCAAGTTTGTGGGTGAGGTGCGTAAAGCACTAAATCGGAACCTTAAAGGAG 300
DB 3397 CTAATCAAGTTTGTGGGTGAGGTGCGTAAAGCACTAAATCGGAACCTTAAAGGAG 3338
QY 301 CCCCCGATTAGACTTGAACGGGGAAGACCGCCGAACGTGGCGAAGAAAGGAGAA 360
DB 3337 CCCCCGATTAGACTTGAACGGGGAAGACCGCCGAACGTGGCGAAGAAAGGAGAA 3278
QY 361 AGCGAAAGAGAGCGGCGCTTAGGCGCTTGCAAGTGTAGCGGTCAACGCTGCGCTAACAC 420
DB 3277 AGCGAAAGAGAGCGGCGCTTAGGCGCTTGCAAGTGTAGCGGTCAACGCTGCGCTAACAC 3218
QY 421 CACACCGCGCGCTTAAATGCGCGCTAAGAGGCGCGTCCCATTCGCCATTCAAGGCTGCG 480
DB 3217 CACACCGCGCGCTTAAATGCGCGCTAAGAGGCGCGTCCCATTCGCCATTCAAGGCTGCG 3158
QY 481 CAACCTGTGGGAAGGCGATGCTGCGCGCTTCTTCCTATTAGCCAGCTGGGAAAGG 540
DB 3157 CAACCTGTGGGAAGGCGATGCTGCGCGCTTCTTCCTATTAGCCAGCTGGGAAAGG 3098
QY 541 GGGATGTGCTCAAGAGCGATTAAATGAGGTAAAGCGAGGTTTCCAGTCAAGCGTTG 600
DB 3097 GGGATGTGCTCAAGAGCGATTAAATGAGGTAAAGCGAGGTTTCCAGTCAAGCGTTG 3038
QY 601 TAAACAGACGCGCAGTGAAGCGCGCTGCTTCAATTCAGTTTTTGAACCCGCTGAGAGAG 660
DB 3037 TAAACAGACGCGCAGTGAAGCGCGCTGCTTCAATTCAGTTTTTGAAGGAGCTGCA 2978
QY 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTATGAGAGCAAGTGAAGTGTGACAC 720
DB 2977 CGCGGTGCGG-----CGCTTGAAGCTAAGTGTGAGTCCCG--CGG 2939
QY 721 GCTCAGAAACAGCAGCTAGATTAACTTGAAGATATATATTTGAGCTAGCTTAA 780
DB 2938 GCTCAGAAATCTTATTTTGTATAGTTTATCATCCATGTGTATATCCAGACAGCTTAA 2879
QY 781 AGATATATCATGCTTAAATTGAGCAGTGAATCTGTATATGACTCATATAGGCGAAT 840
DB 2878 CAAACTCAAGAAAGACCATGTGTACGCTTTTGTGTGAGTCTTTGAAAGGCGAAGTT 2819

QY 841 TGGGTACCGGCGCCCGCTCGAGGTGACAGGTATCGATTAAGCTTATATGAAATTCCTCG 900
DB 2818 GTGTGACAGGTAAATGTTGTCTGTAAAGAGAGAGGCGCATCGCCAAATGAGATATTT 2759
QY 901 AGCCCGGCGGATTCAGTATCTAGAGCGGCGCCAGCGGCTGAGCTCCAGCTTTTGT 960
DB 2758 GTTATATATGCTCTGTAGTGAACGATTCATCTTCAATGTGTGCGAATTTTGAAGT 2699
QY 961 TCCCTTATGAGGCTTAATTAATGATCCAGTGTCAATTTTGAAGCTATCTTTCTAG 1020
DB 2698 TAGCTTATGATTCATCTTTTGTGTGTGCTGCGGATGATATCATGTGTAG--TTATAG 2640
QY 1021 GGTATATGATGCTCATCAGATATCATTCGCGGCTTTTTCGCGCTCATGCTATCGCC 1080
DB 2639 TTGATCTGAGTGTGTGCTCGAGATGTTCATCTTTTAAATCAATACCTTTTAAAC 2580
QY 1081 CAACCTGCGCTATCTGCGCATTCGAGGAGAGAAAGCCCGTCTTTTCCCGCAGGTTG 1140
DB 2579 TCGATAGGATTAACAAGGATATCACCTTCAAACTTGACTGACAGCGCTTGTAGTTG 2520
QY 1141 AAGCGCATGGAAGAGTTTCCGAGATGACTGCTGCTGATTAAGCTTGAGCGAAAC 1200
DB 2519 CCGTCATCTTTGAAGATATAGTGTCTGCTGTATCATTAACCTTGGCGCATGCACTTGG 2460
QY 1201 GCAGCTTACCATGATATTCGGAAGTGTGCGCATGCAACGCTTTAAACGCTGAACCTGT 1260
DB 2459 AAAAAGTATGCGCTTCAATATGATCCGATTAAGGAGAAAGCAATGAACCATTAAGAG 2400
QY 1261 TCGTTACGCGCACCTGAGATACAGTTCGCGCTTTTCCGACACAGTTCCGATGG 1320
DB 2399 AAATATGATGACATGTGTGCGCATGGAACAGTATGTTTCCAGTATGTCAAATTAATTA 2340
QY 1321 TCAGCCGGAAGCGCATGACAAACCGAACAATACCGCGACAGCGGAACTGCGTCCG 1380
DB 2339 AGGATAGCTTTCCGATGTAGCATCACTTCAACCTCTCACTGACAGAAATTTGTGC 2280
QY 1381 GTGTGAGATTAATGACAGCGGTGCGCGCTGGAATTTACGTACGAGAGAGACGGATATC 1440
DB 2279 CCAATTAATCATCACCATTAAATTCAAAGAAATTTGGAACAACTCCAGTAAAGTTCTCT 2220
QY 1441 CTGCGCTGATGCGCAGAAATGACATGATTAACCCCTGAGTTACCGGCGGCGCGCTT 1500
DB 2219 C---CTTATCTCATCGGTATCCAGCTTTTGTCTCTTATGAGAGGTTAATGTGCGCTT 2163
QY 1501 GCGGTAAATCATGTGATAGCTGTTTCTGTGTGAATTTGATTCGCTCAAAATTCACA 1560
DB 2162 GCGGTAAATCATGTGATAGCTGTTTCTGTGTGAATTTGATTCGCTCAAAATTCACA 2103
QY 1561 CAACATACAGAGCGGAAAGCATTAAGTAAAGCTGAGGTGCTTAATGATGAGCTTAAT 1620
DB 2102 CAACATACAGAGCGGAAAGCATTAAGTAAAGCTGAGGTGCTTAATGATGAGCTTAAT 2043
QY 1621 CACATTAATGCGGTGCTCACTGCGCGCTTCACTCGGGAACCTGTCGTGCTCAAGCT 1680
DB 2042 CACATTAATGCGGTGCTCACTGCGCGCTTCACTCGGGAACCTGTCGTGCTCAAGCT 1963
QY 1681 GCATTAATGATGCGGCAACGCGGAGAGAGCGGTTTCGTAATGAGCGCTTTCGCG 1740
DB 1982 GCATTAATGATGCGGCAACGCGGAGAGAGCGGTTTCGTAATGAGCGCTTTCGCG 1923
QY 1741 TTCTGTGCTCACTGACCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 1922 TTCTGTGCTCACTGACCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
QY 1801 CTCAAAAGCGGTAAATCGTTATTCACAGAAATCAGGAGATTAAGCAGAAAGAAACATGTG 1860
DB 1862 CTCAAAAGCGGTAAATCGTTATTCACAGAAATCAGGAGATTAAGCAGAAAGAAACATGTG 1803
QY 1861 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACGCTTAAGAGCGGCTTGTGCTGCTTTTCCA 1920
DB 1802 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACGCTTAAGAGCGGCTTGTGCTGCTTTTCCA 1743
QY 1921 TAGGCTCCGCGCCCTGACAGCATCACAAAATTCAGCTCAAGTCAAGGTGCGCAAA 1980

```

Db      1742 TAGGCTCCGCCCCCTGACGAGCATCAAAAATTCAGCTCAAGTCAGAGGTGGGAAA 1683
Qy      1981 CCCGACAGGACTATTAAGATACACAGCGTTTCCCGCTGGAAGTCCCTGTCGCTCC 2040
Db      1682 CCCGACAGGACTATTAAGATACACAGCGTTTCCCGCTGGAAGTCCCTGTCGCTCC 1623
Qy      2041 TGTTCGACCGCTGCGGCTTACCGGATACCTGTCGCTTTCCTCCGCGGAGCGTGC 2100
Db      1622 TGTTCGACCGCTGCGGCTTACCGGATACCTGTCGCTTTCCTCCGCGGAGCGTGC 1563
Qy      2101 GCTTTCATAGCTCAAGCTGTAGATCTCAGTTCCGTTAGGTCGTTCCCTCAAGCT 2160
Db      1562 GCTTTCATAGCTCAAGCTGTAGATCTCAGTTCCGTTAGGTCGTTCCCTCAAGCT 1503
Qy      2161 GGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGTAACATA 2220
Db      1502 GGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGTAACATA 1443
Qy      2221 TCTTAGATCCAAACCGGTAAACAGACTATCCGCTACGCGAGCCACTGTAAAC 2280
Db      1442 TCTTAGATCCAAACCGGTAAACAGACTATCCGCTACGCGAGCCACTGTAAAC 1383
Qy      2281 GATTAGCAGAGCAGAGTATGTAGCGGTCTACAGAGTCTTGAAGTGTGGCTAACTA 2340
Db      1382 GATTAGCAGAGCAGAGTATGTAGCGGTCTACAGAGTCTTGAAGTGTGGCTAACTA 1323
Qy      2341 CGGCTACACTGAAGAGCAGTATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db      1322 CGGCTACACTGAAGAGCAGTATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263
Qy      2401 AAAAAGTGTGTAGCTCTTATCGGCAAAACCAACCGCTGTAGCGGTGTTTTT 2460
Db      1262 AAAAAGTGTGTAGCTCTTATCGGCAAAACCAACCGCTGTAGCGGTGTTTTT 1203
Qy      2461 TGTTCGAACACAGATTAACGCGGAGAAAAGATCTCAAGAAATCCTTTGATCTT 2520
Db      1202 TGTTCGAACACAGATTAACGCGGAGAAAAGATCTCAAGAAATCCTTTGATCTT 1143
Qy      2521 TTTCACGGGCTGACGCTCAGTGAAGAAAACCTCAGCTTAAGGATTTTGTCAATG 2580
Db      1142 TTTCACGGGCTGACGCTCAGTGAAGAAAACCTCAGCTTAAGGATTTTGTCAATG 1083
Qy      2581 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAAT 2640
Db      1082 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAAT 1023
Qy      2641 CTAAAGTATATATGAATTAACCTGATCTGACAGTTACATGCTTATAGAGGAC 2700
Db      1022 CTAAAGTATATATGAATTAACCTGATCTGACAGTTACATGCTTATAGAGGAC 963
Qy      2701 TATCTCAGCAGATCTGATCTTATCTTATCATATGTTGCTGACTCCCGTGTATAT 2760
Db      962 TATCTCAGCAGATCTGATCTTATCTTATCATATGTTGCTGACTCCCGTGTATAT 903
Qy      2761 AACTACGATACGGAGGCTTACATCTGACCGCACTGCTGCAATGATACCGGAGACC 2820
Db      902 AACTACGATACGGAGGCTTACATCTGACCGCACTGCTGCAATGATACCGGAGACC 843
Qy      2821 ACGCTCACCGGCTCCAGATTATACAGATTAACCAAGCAGCGGAGGCGGAGCCAG 2880
Db      842 ACGCTCACCGGCTCCAGATTATACAGATTAACCAAGCAGCGGAGGCGGAGCCAG 783
Qy      2881 AATGTGCTCTGCAACTTTATCCGCTCATCCAGCTTATATTTGCGGGAGAGTAG 2940
Db      782 AATGTGCTCTGCAACTTTATCCGCTCATCCAGCTTATATTTGCGGGAGAGTAG 723
Qy      2941 AGTAAGTATGTCGCAATTAAGTTTGGCAAGCTTGTGCTACAGGATCTG 3000
Db      722 AGTAAGTATGTCGCAATTAAGTTTGGCAAGCTTGTGCTACAGGATCTG 663
Qy      3001 GGTGTACGCTGTGTGTGTATGCTTATAGCTCCGCTTCCAGATCAAGGCG 3060

```

```

Db      662 GGTGTACGCTGTGTGTGTATGCTTATAGCTTACGCTCCGCTTCCAGATCAAGGCG 603
Qy      3061 AGTTACATGATCCCCATGTTGTGCAAAAAGGCTTATAGTCTTGTGCTCCGATCTG 3120
Db      602 AGTTACATGATCCCCATGTTGTGCAAAAAGGCTTATAGTCTTGTGCTCCGATCTG 543
Qy      3121 TGTGCAAGTATGTTGCGCCGACAGTGTATCACTCATGTTATAGGACAGCTCATATTC 3180
Db      542 TGTGCAAGTATGTTGCGCCGACAGTGTATCACTCATGTTATAGGACAGCTCATATTC 483
Qy      3181 TCTTACTGTATGTCATCCGTAAGATGCTTTCTGTGACTGTGTGACTAACCAGTGC 3240
Db      482 TCTTACTGTATGTCATCCGTAAGATGCTTTCTGTGACTGTGTGACTAACCAGTGC 423
Qy      3241 ATTCTGAGATATGTTATGTCGCGGACCGAGTGTCTTGTCCGCGCTCAATAGGAGTAA 3300
Db      422 ATTCTGAGATATGTTATGTCGCGGACCGAGTGTCTTGTCCGCGCTCAATAGGAGTAA 363
Qy      3301 TACCGGCGCATAGCAGAACTTTAAAGTGTCTATCATTTGAAAAAGCTTCTGCGGCG 3360
Db      362 TACCGGCGCATAGCAGAACTTTAAAGTGTCTATCATTTGAAAAAGCTTCTGCGGCG 303
Qy      3361 AAAACTCTCAAGATCTTACCGCTGTGTGATCCAGTTGATTAACCACTGTGACCC 3420
Db      302 AAAACTCTCAAGATCTTACCGCTGTGTGATCCAGTTGATTAACCACTGTGACCC 243
Qy      3421 CAACTGATCTTCAAGATCTTATCTTACCAAGGCTTCTGCGTGTGAGCAAAAACGGAAG 3480
Db      242 CAACTGATCTTCAAGATCTTATCTTATCTTACCAAGGCTTCTGCGTGTGAGCAAAAACGGAAG 183
Qy      3481 GCAAAATGCGCAAAAAGGAAATAGGGGACAGCGAAATGTGAATCTGATCTCTT 3540
Db      182 GCAAAATGCGCAAAAAGGAAATAGGGGACAGCGAAATGTGAATCTGATCTCTT 123
Qy      3541 CCTTTTCATATATTAAGAACTTATACAGGCTTATGTCATAGACGATACATAT 3600
Db      122 CCTTTTCATATATTAAGAACTTATACAGGCTTATGTCATAGACGATACATAT 63
Qy      3601 TGAATGATTTTGAAGAAATTAACAAATAGGGCTTCCGCGCAATTTCCCGAAAGTGC 3660
Db      62 TGAATGATTTTGAAGAAATTAACAAATAGGGCTTCCGCGCAATTTCCCGAAAGTGC 3
Qy      3661 AC 3662
Db      2 AC 1

```

RESULT 12
 US-10-280-913A-17/c
 / Sequence 17, Application US/10280913A
 / Publication No. US20040110130A1
 / GENERAL INFORMATION:
 / APPLICANT: Large Scale Biology Corporation
 / TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
 / FILE REFERENCE: LSBC-PLG5392-CIP
 / CURRENT APPLICATION NUMBER: US/10/280, 913A
 / PRIOR FILING DATE: 2002-10-25
 / PRIOR APPLICATION NUMBER: 60/402,342
 / PRIOR FILING DATE: 2002-08-08
 / PRIOR APPLICATION NUMBER: 10/066,390
 / PRIOR FILING DATE: 2002-02-01
 / PRIOR APPLICATION NUMBER: 60/268,785
 / PRIOR FILING DATE: 2001-02-14
 / PRIOR APPLICATION NUMBER: 60/266,386
 / PRIOR FILING DATE: 2002-02-02
 / NUMBER OF SEQ ID NOS: 35
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 17
 / LENGTH: 3637
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Encodes cycle 3 GPP

US-10-280-913A-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;
Bee Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

1 CTAATTTGTAAGCGTTAATATTTTGTAAATTTGCGTTAAATTTTGTAAATGACGTC 60
|||
3637 CTAATTTGTAAGCGTTAATATTTTGTAAATTTGCGTTAAATTTTGTAAATGACGTC 3578
61 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTAATAATCAAAAGATAGACGA 120
3577 ATTTTAAACCAATAGCGCGAAATCGCAAAATCCCTTAATAATCAAAAGATAGACGA 3518
121 GATAGGGTTGAGTGTGTTTCCAGTTTGAACAAGATCCATATTAAAGAACTGAGACTC 180
3517 GATAGGGTTGAGTGTGTTTCCAGTTTGAACAAGATCCATATTAAAGAACTGAGACTC 3458
181 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGCGCCACTAGCAATCAATCAC 240
3457 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGCGCCACTAGCAATCAATCAC 3398
241 CTAATCAAGTTTGTGGGTGAGGTGCGCTAAAGCACTAAATCGAAACCTTAAAGGAG 300
3397 CTAATCAAGTTTGTGGGTGAGGTGCGCTAAAGCACTAAATCGAAACCTTAAAGGAG 3338
301 CCCCCGATTGAGCTTGAACGGGGAAAGCCGCGCAACGTGCGGAGAAAGAGAGAA 360
3337 CCCCCGATTGAGCTTGAACGGGGAAAGCCGCGCAACGTGCGGAGAAAGAGAGAA 3278
361 AGCGAAAGAGCGGGCGCTAGGAGCGCTGCAAGTGTAGCGGTCAAGCGCGGTAAACAC 420
3277 AGCGAAAGAGCGGGCGCTAGGAGCGCTGCAAGTGTAGCGGTCAAGCGCGGTAAACAC 3218
421 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCATTCGCCATTCAAGCTCG 480
3217 CACACCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCATTCGCCATTCAAGCTCG 3158
481 CAACCTGTGGGAAGGGCGATGCTGCGGGCTCTTCCGTATTAACCGCAGCTGCGAAAG 540
3157 CAACCTGTGGGAAGGGCGATGCTGCGGGCTCTTCCGTATTAACCGCAGCTGCGAAAG 3098
541 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAACGCAAGGTTTTCCCGACGACGATG 600
3097 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAACGCAAGGTTTTCCCGACGATG 3038
601 TAAACGACGCGCAGTGAACGCGCTCGTTCATTCACTTTTGAACCGGTGAGAGAG 660
3037 TAAACGACGCGCAGTGAACGCGCTCGTTCATTCACTTTTGAACCGGTGAGAGAG 2978
661 GCAAGCTGCGCGTGAATGTGTTTACAGCGTGTGAGAGATGAAGATGCTGACAC 720
2977 CCGCGGTGCGCG-----CCGCTAGAACTAGTGAATCCG--CGG 2939
721 GCTCAGAACACGACGCTAGATTAACCTAGAAAGTAAATCAATGTGAGAGTACGTTAA 780
2938 GCTCAGAACATTTCTTAATGTAGTATTCATCAGTCAGTGTATTCACACAGAGTTA 2879
781 AGATATTCATCGCTAAATTTGACGATGGATCTGTATTAAGACTCACTATAGGCGGAT 840
2878 CAATCTCAAGAAAGGACATGTGTGACGCTTTTGTGGGATCTTTGCAAAAGGCAAGTT 2819
841 TGGGTACCGGGCCCCCTCGAGGTGACGGGTATGATAGCTTGATATTCGAATTCCTGC 900
2818 GTGTGACAGGTAAATGTGTGTCTGTGTAAGAACAGCAGGCGCATTCGCAATTTGAGATTTT 2759
901 AGCCGCGGAGATTCACCTAGTCTAGAGGCGCGCAACGCGGTGAGAGTTCAGCTTTTGT 960
2758 GTTATATATGCTGTGTAGTGTGAACGATTCATCTTCAATGTGTGGGAAATTTGAAGT 2699
961 TCCCTTATAGTAGGATTAATTAGATCCCATGCTCAATTTTACGACGACTATCTTTCTAG 1020
2698 TAGCTTTATTCATTTCTTTTGTGTGTGCGCGGTATGTATCAATTTGTGTAG--TTATAG 2640

1021 GGTATATCTACGTGATCAGATCATATCGTGGGCTCTTTTTCGGCTCAGTATCGCC 1080
2639 TTGTATCTGAGATTTGTGTGCTTCCAGAAATGTTTCCATCTTTTAAATCAATCTTTTAAAC 2580
1081 CAAGCTGCGCTATCTGGGCGATCGGGGAGAAAGCCCGGCTTTTCCGCGAGGTTG 1140
2579 TCGATAGCATTAACAAGGATATCACCTTCAACTTGAACCTTACAGCAGCGCTTTGTAGTTC 2520
1141 AAGCGGATGAAAGAGTTTCCGAGATGACCTGCTGCTGATGACGTTGAGCGAAAC 1200
2519 CCGTATCTTTGAAAGATATGAGCGTTCCGTATCATTAACCTTGGGCGATGACCTGTG 2460
1201 GCACGTTTACATATGATATCGGAGAGTGTGCGCATGACGCGCTTTAAGCGTGAATGT 1260
2459 AAAAAGTATCCCTTTTCAATATGATCCGGAATACGGGAAAGCATTTGAACCATATAGAG 2400
1261 TCGTTCAAGGCGACCGGGAATACAGATTGCTGTGGGCTTTTCGGACACAGTTCCGGATG 1320
2399 AAAGTATGACAGATGTTGGCCATGAAACAGTATGTTTCCAGTATGCAATTAATTTA 2340
1321 TCAGCCCGAAGCGCATGACAAACCGGAATAATCCGCGACAGCGGAACTGCGGTGCG 1380
2339 AGGGTATAGCTTCCGTATGTATGATCACTTCAACCTCTCCATGACAGAAATTTGTGC 2280
1381 GTGTGCAATTAATGACAGCGGTGCGGCTGGGATTTATAGTCAAGCGAGAGCGGATTC 1440
2279 CCATTTAATCATCACCATCTAATTAACAAGATTTGGGCAACTCCAGTGAATAAGTCTTCT 2220
1441 CTGGCTGGAATCGCGCAAAATGSAATGSAATACCCGGAAGTTACCGGCGGGGCGGCTT 1500
2219 C---CTTTTCTCATCGGTAACCACTTTTGTCTTTTGTAGAGGTTAATTTGCGCGCTT 2163
1501 GCGCTAATCATGATGATAGCTGTTTCTGTGTGAATTTGTATCCGCTCACAATTTCA 1560
2162 GCGGTATCATGATGATAGCTGTTTCTGTGTGAATTTGTATCCGCTCACAATTTCA 2103
1561 CAACATACAGCGCGAAAGCATTAAGTTAAAGCTTGGGCTGCTTAATGATGATTAAT 1620
2102 CAACATACAGCGCGAAAGCATTAAGTTAAAGCTTGGGCTGCTTAATGATGATTAAT 2043
1621 CACATTAATGCGTGTGCGCTACGCGCGCTTCCAGTGGGAAACGTTGTGTCAGACT 1680
2042 CACATTAATGCGTGTGCGCTACGCGCGCTTCCAGTGGGAAACGTTGTGTCAGACT 1983
1681 GCATTTAATGATCGGCAACGCGGGGAGAGCGGTTTGTGATTTGGGCGCTCTTCCG 1923
1982 GCATTTAATGATCGGCAACGCGGGGAGAGCGGTTTGTGATTTGGGCGCTCTTCCG 1923
1741 TTCTCTGCTCATGATCTGCTGCGCTGTGTGCTTGGCTGTGCGAGCGGTATCACTCA 1800
1922 TTCTCTGCTCATGATCTGCTGCGCTGTGTGCTTGGCTGTGCGAGCGGTATCACTCA 1863
1801 CTCAAGCGGTATATGAGTTATCCAGAAATCAGAGATGAGGAGTAAACGAGAAACATGTG 1860
1862 CTCAAGCGGTATATGAGTTATCCAGAAATCAGAGATGAGGAGTAAACGAGAAACATGTG 1803
1861 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACGTTAAAGGCGCGCTTGTGCGGTTTTC 1920
1802 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACGTTAAAGGCGCGCTTGTGCGGTTTTC 1743
1921 TAGGCTCGCGCCCCCTGATGACGATCACAAAATTCAGCTCAAGTCAAGAGTGTGCA 1980
1742 TAGGCTCGCGCCCCCTGATGACGATCACAAAATTCAGCTCAAGTCAAGAGTGTGCA 1683
1981 CCGGACAGGATTAATTAAGATACAGAGGCTTTCCCGCTGGAAGCTCCCTGCGGCTCTCC 2040
1682 CCGGACAGGATTAATTAAGATACAGAGGCTTTCCCGCTGGAAGCTCCCTGCGGCTCTCC 1623
2041 TGTTCGACCTGTGCGCTTACCGGATCTGCTGCGCTTCTCCCTTTCGGGAAAGCGTGGC 2100
1622 TGTTCGACCTGTGCGCTTACCGGATCTGCTGCGCTTCTCCCTTTCGGGAAAGCGTGGC 1563

QY 2101 GCTTCTCATAGTCAAGCTGTAAGTATCTAGTTCGATGAGGAGGCTTCAAGCT 2160
 Db 1562 GCTTCTCATAGTCAAGCTGTAAGTATCTAGTTCGATGAGGAGGCTTCAAGCT 2160
 QY 2161 GGGCTGTGTGCAAGAACCCCGTTCACCGGACCGGCTTATCCGTAACATATG 1503
 Db 1502 GGGCTGTGTGCAAGAACCCCGTTCACCGGACCGGCTTATCCGTAACATATG 1503
 QY 2221 TCTTGAATCAACCCGGTAAGACAGATTAATGCGCACTGGCAGCCATGTGTAACG 1443
 Db 1442 TCTTGAATCAACCCGGTAAGACAGATTAATGCGCACTGGCAGCCATGTGTAACG 1443
 QY 2281 GATTAGCAGACGAGGATGTAGGCGGTGCTACAGAGTCTTGAATGCTGCTTAATCTA 1383
 Db 1382 GATTAGCAGACGAGGATGTAGGCGGTGCTACAGAGTCTTGAATGCTGCTTAATCTA 1383
 QY 2341 CGGCTACACTAAGAGACAGATTAATGCTGCTGCTGCTGTAAGCCAGTTACCTTGG 2400
 Db 1322 CGGCTACACTAAGAGACAGATTAATGCTGCTGCTGCTGTAAGCCAGTTACCTTGG 1263
 QY 2401 AAAAAGATTGTAAGTCTTGAATCCGGCAACAAACACCGCTGCTGCTGCTTCTTCTT 2460
 Db 1262 AAAAAGATTGTAAGTCTTGAATCCGGCAACAAACACCGCTGCTGCTGCTTCTTCTT 1203
 QY 2461 TGTTCGACAGCAGCAGATTAACCGCAGAAAGAAAGATCTCAGAGAGTCTTGAATCTT 2520
 Db 1202 TGTTCGACAGCAGCAGATTAACCGCAGAAAGAAAGATCTCAGAGAGTCTTGAATCTT 1143
 QY 2521 TTTCAAGGGGTCTGAGCTCAGTGGAAAGAAAGATCTCAGAGAGTCTTGAATCTT 2580
 Db 1142 TTTCAAGGGGTCTGAGCTCAGTGGAAAGAAAGATCTCAGAGAGTCTTGAATCTT 1083
 QY 2581 ATTATCAAAAGAGATCTTCACTGATCTTTTAAATTAATGAAGTTTAAATCAAT 2640
 Db 1082 ATTATCAAAAGAGATCTTCACTGATCTTTTAAATTAATGAAGTTTAAATCAAT 1023
 QY 2641 CTAAGATATATAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 Db 1022 CTAAGATATATAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2701
 QY 2701 TATCTCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
 Db 2761 TATCTCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2761
 QY 2761 AACTACGATACGGAGGGCTTACCACTGCGCCCAAGCTGCAATGATATACCGGACACCC 2820
 Db 902 AACTACGATACGGAGGGCTTACCACTGCGCCCAAGCTGCAATGATATACCGGACACCC 2821
 QY 2821 ACGCTACCGGCTCCAGATTTATCACAATTAACCAAGCCAGCGGAGGCGGAGCGCAG 2880
 Db 842 ACGCTACCGGCTCCAGATTTATCACAATTAACCAAGCCAGCGGAGGCGGAGCGCAG 2881
 QY 2881 AAGTGTCTCTGCAATTTATCCGCTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
 Db 782 AAGTGTCTCTGCAATTTATCCGCTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2941
 QY 2941 AGTAAGTATGCGCACTTAATAGTTTGGCAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Db 722 AGTAAGTATGCGCACTTAATAGTTTGGCAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 3001
 QY 3001 GGTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
 Db 662 GGTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3061
 QY 3061 AGTTACATGATCCCGCATGTGTGTGCAAAAGCGGTAGCTCTTGGCTCTCGATCGT 3120
 Db 602 AGTTACATGATCCCGCATGTGTGTGCAAAAGCGGTAGCTCTTGGCTCTCGATCGT 3121
 QY 3121 TGTCAAAATTAAGTTGGCGCAGTGTATCACTGATGTTATGCGACATGCTGATATTC 3180
 Db 542 TGTCAAAATTAAGTTGGCGCAGTGTATCACTGATGTTATGCGACATGCTGATATTC 3181
 QY 3181 TCTTACTGTATGATCCGTAAGTGTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240

Db 482 TCTTACTGTATGATCCATCCGTAAGTGTCTTCTGTAACGTAAGTACTCAACAGTTC 423
 QY 3241 ATTCTGAGATATGATGTATGCGGCGACGAGTGTCTTGGCGGCGCAATACGGATTA 3300
 Db 422 ATTCTGAGATATGATGTATGCGGCGACGAGTGTCTTGGCGGCGCAATACGGATTA 3301
 QY 3301 TACCGGCGCAATGACAGAACTTTAAAGTGTCTATGTAAGAAAGCTTCTTGGGCG 3360
 Db 362 TACCGGCGCAATGACAGAACTTTAAAGTGTCTATGTAAGAAAGCTTCTTGGGCG 303
 QY 3361 AAACTCTCAAGATCTTACCGGCTGTTGAGTTCAGTGTGATGTAACCACTGCTGAC 3420
 Db 302 AAACTCTCAAGATCTTACCGGCTGTTGAGTTCAGTGTGATGTAACCACTGCTGAC 243
 QY 3421 CAACTGATCTTCTGCAATCTTCTTCACTTCAACGAGCTTCTGAGTACGAAACAGGAG 3480
 Db 242 CAACTGATCTTCTGCAATCTTCTTCACTTCAACGAGCTTCTGAGTACGAAACAGGAG 183
 QY 3481 GCAAAATGCGCAAAAGGAAATAGGCGCAACGAAATGTTGATCTCATCTCTT 3540
 Db 182 GCAAAATGCGCAAAAGGAAATAGGCGCAACGAAATGTTGATCTCATCTCTT 123
 QY 3541 CCTTTTCAATTTATGAGCAATTTATCAGGCTTATGCTCATAGAGGATATCTT 3600
 Db 122 CCTTTTCAATTTATGAGCAATTTATCAGGCTTATGCTCATAGAGGATATCTT 63
 QY 3601 TGAATGTATTTGAAATTAACAAATAGGAGTTCGCGCACTTCCCGAAAGTCC 3660
 Db 62 TGAATGTATTTGAAATTAACAAATAGGAGTTCGCGCACTTCCCGAAAGTCC 3
 QY 3661 AC 3662
 Db 2 AC 1

RESULT 13
 US-10-684-134-17/c
 ; Sequence 17, Application us/10684134
 ; Publication No. US2004014243A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Large Scale Biology Corporation
 ; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
 ; FILE REFERENCE: LSBC-PLG5392-CIP
 ; CURRENT APPLICATION NUMBER: US/10/684,134
 ; PRIOR FILING DATE: 2003-10-10
 ; PRIOR APPLICATION NUMBER: 60/402,342
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 10/066,390
 ; PRIOR FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: 60/268,785
 ; PRIOR FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: 60/266,386
 ; PRIOR FILING DATE: 2002-02-02
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 17
 ; LENGTH: 3637
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURES:
 ; OTHER INFORMATION: Encodes cycle 3 GFP
 US-10-684-134-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;
 Best Local Similarity 83.8%; Pred. No. 0;
 Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;
 QY 1 CTAATTTGAGCGTATATTTTGTAAATTCGCTTAATTTTGTAAATCACTC 60
 Db 3637 CTAATTTGAGCGTATATTTTGTAAATTCGCTTAATTTTGTAAATCACTC 60
 QY 61 ATTTTAAACCAATAGCCGCAAAATCCGCAAAATCCCTTATTAATCAAAAGATGACCGA 120

Db 3577 ATTTTAAACCAATAGCCGAAATCGCAAAATCCCTTATTAATCAAAAGAAATAGACCA 3518
Qy 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTATTAAAGACGTGACTC 180
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTATTAAAGACGTGACTC 3458
Qy 181 CAACGTCAAAAGGCGCAAAACCGTCTATCAAGGCGCATGGCCCACTAGCTGAACCTACCC 240
Db 3457 CAACGTCAAAAGGCGCAAAACCGTCTATCAAGGCGCATGGCCCACTAGCTGAACCTACCC 3398
Qy 241 CTAATCAAGTTTGTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAGGGAG 300
Db 3397 CTAATCAAGTTTGTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAGGGAG 3338
Qy 301 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGGCGAACCTGGCGAAGAAAGAAAGGAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGGCGAACCTGGCGAAGAAAGGAGAAAG 3278
Qy 361 AGCGAAAGAGACGGGCGCTAGAGGCGCTGAGAGTACCGGTCAAGCTGCGCGTAAACAC 420
Db 3277 AGCGAAAGAGACGGGCGCTAGAGGCGCTGAGAGTACCGGTCAAGCTGCGCGTAAACAC 3218
Qy 421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAAGCTCG 480
Db 3217 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAAGCTCG 3158
Qy 481 CAACGTGTGGAAAGGCGATCGGTGCGGCGCTTCTTGCTATTAACGCACTGGCGAAGG 540
Db 3157 CAACGTGTGGAAAGGCGATCGGTGCGGCGCTTCTTGCTATTAACGCACTGGCGAAGG 3098
Qy 541 GGGATGCTGCTCAAGGCGATTAAGTGGGTAAACGCAAGGTTTTCCAGTCAACGCTTG 600
Db 3097 GGGATGCTGCTCAAGGCGATTAAGTGGGTAAACGCAAGGTTTTCCAGTCAACGCTTG 3038
Qy 601 TAAACGACGCGCAAGTGAAGCGCGCTCGTTCACTTCAAGTTTTGAACCGGAGAGAGG 660
Db 3037 TAAACGACGCGCAAGTGAAGCGCGCTCGTTCACTTCAAGTTTTGAACCGGAGAGG 2978
Qy 661 GCAAGCTCGCGGTGCAAAATGTGTTTTACAGCGTGAATGAGCAATGAAGTCTGACAC 720
Db 2977 CCGCGGTGCGG-----CGGCTTGAAGATGATGATCCG--CGG 2939
Qy 721 GCTCGAAGACACGACGTAATGAATCACTTAAGAAATATCATATTGACGATGTTAA 780
Db 2938 GCTCGAAGAAATTTATTTATGATGATTCATCATGTCATGATATCCACAGACAGTTA 2879
Qy 781 AGATAATCATGCTGAATAATTGACGATGGATCTGTAAATAGACCTCACTATAGGCGAAT 840
Db 2878 CAACCTCAAGAAAGACATGTGTGACGCTTTTGTGGGATCTTTGCAAAAGGCGAAT 2819
Qy 841 TGGGTACGGGCGCCCCCTGAGGTGACGATGATGATGATGATGATGATGATGATGAT 900
Db 2818 GTGTGACAGGTAATGTGTCTGGTAAAGAACAGGCGCATTCGCAATTTGAGATTTT 2759
Qy 901 AGCCCCGGGGATTCACATGTTCTAGAGCGCGCCACCGCGTGAAGTCCAGCTTTGT 960
Db 2758 GTTGAATAATGCTGTGATGTAAGAACGATTCATCTTCAATGTGTGGGAAATTTGAGT 2699
Qy 961 TCCCTTAGTGAAGGTTAATTAAGATCCCATGCTCAATTTTACGACATATCTTTAG 1020
Db 2698 TAGCTTTGATTCATCTTTTGTGTTGTGTCGCGTGAATATATGATGATGATGATGATG 2640
Qy 1021 GATTATCTAGCTGATCAGATCAATATGATGATGATGATGATGATGATGATGATGATG 1080
Db 2639 TTGTACTGAGTTGTGTGCGAAGATGTTTCCATCTTCTTAAATCAATATCTTTTAAAC 2580
Qy 1081 CAACCTGGCGCTATCTGGGCAATCGGGAAGAAAGCCGCTGCTTTTCCGCGACATGCGC 1140
Db 2579 TCGATACGATTAACAAGGATACATTTCAAACTTTGATTCAGACGCGCTTTGTAGTTC 2520
Qy 1141 AAGCGCATGGAAGAGTTTCCGAGATGACGTGCTGATGACGTTGAGGAAAC 1200
Db 2519 CCGTCACTTTGAAGATATAGTGCTTCTGTATCATTAACCTTGGGATGGAACCTCTTG 2460
Qy 1201 GCACGTTTACATGATGATTTGGGAAGGTGTGGCATGACAGCTTTAAAGGTGAACGTG 1260
Db 2459 AAAAAGTCATCCGTTTATATGATCCGATTAACGGGAAAGATTAAGAACATTAAGAG 2400
Qy 1261 TCGTTCAAGCCACCTGGGATACAGTGTGTGCGGCTTTTCCGACACATGTCGGATG 1320
Db 2399 AAAGTATGAACAGTTGGTGGCATGGAACAGTGTGTTTTCCAGTATGTAATTAATTTA 2340
Qy 1321 TCAGCCCGAAGCGATCAGCAACCGAACAATTAACCGGCGACAGCGGAACTCGGTGCG 1380
Db 2339 AGGTGAAGCTTCCGATATGATGATCACTTCAACCTCTCCATCAGCAAGAAATTTGTGC 2280
Qy 1381 GTGTGACATTAATGACAGCGGTGCGGCTGGGATTAATTAAGTACGCGAGACGGGTATC 1440
Db 2279 CCATTAACATCAACATCTTAATTAACAAGAAATTTGGACAACTCCAGTGAATAATTTCT 2220
Qy 1441 CTGCTGATGTCGCCGCAAAATGACATGGAATACCCCGTGAATACCGGCGGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTCTTTAGTGAAGGTTAATTTGGCGCTT 2163
Qy 1501 GCGGTAATCATGATCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCAATTTCCACA 1560
Db 2162 GCGGTAATCATGATCATAGCTGTTTCTGTGTGAATTTGTAATCGCTCAATTTCCACA 2103
Qy 1561 CAACATACAGACCGGAAGCATTAAGTAAAGCTGGGGTGCCTTAATGATGACTAACT 1620
Db 2102 CAACATACAGACCGGAAGCATTAAGTAAAGCTGGGGTGCCTTAATGATGACTAACT 2043
Qy 1621 CACATTAATGCGTTGGGCTCACTGCGCTTTTCACTCGGAAACCTGTGTGCTGACT 1680
Db 2042 CACATTAATGCGTTGGGCTCACTGCGCTTTTCACTCGGAAACCTGTGTGCTGACT 1983
Qy 1681 GCATTAATGAAATCGGCAACGCGCGGGAAGAGCGTTTGTGTAATTTGGGCGCTCTCCG 1740
Db 1982 GCATTAATGAAATCGGCAACGCGCGGGAAGAGCGTTTGTGTAATTTGGGCGCTCTCCG 1923
Qy 1741 TTCTGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1922 TTCTGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
Qy 1801 CTCAAAAGCGGTAATACGTTATTCACAGATCAGGGGATTAACGCAAGAAAGACATGTG 1860
Db 1862 CTCAAAAGCGGTAATACGTTATTCACAGATCAGGGGATTAACGCAAGAAAGACATGTG 1803
Qy 1861 AGCAAAAGGCGCAGCAAAAGGCAAGAACCGTAATAAAGCGCGCTGTGGGCTTTTCCA 1920
Db 1802 AGCAAAAGGCGCAGCAAAAGGCAAGAACCGTAATAAAGCGCGCTGTGGGCTTTTCCA 1743
Qy 1921 TAGGCTCGGCCCCCTGACGATCACAAAATCGACGCTCAAGTCAAGAGGTGCGAA 1980
Db 1742 TAGGCTCGGCCCCCTGACGATCACAAAATCGACGCTCAAGTCAAGAGGTGCGAA 1683
Qy 1981 CCGCAACGACATTAAGATTAACAGAGGCTTTCCCTCTGAAAGCTCTCTGCTCTCC 2040
Db 1682 CCGCAACGACATTAAGATTAACAGAGGCTTTCCCTCTGAAAGCTCTCTGCTCTCC 1623
Qy 2041 TGTTCGACACCTGCGCTTACCGGATACGCTGTGCGCTTTCTCCCTCGGGAAGCGTGGC 2100
Db 1622 TGTTCGACACCTGCGCTTACCGGATACGCTGTGCGCTTTCTCCCTCGGGAAGCGTGGC 1563
Qy 2101 GCTTTCATAGCTCAAGCTGTAGATCTCAAGTTCCGATGATGATGATGATGATGATGATG 2160
Db 1562 GCTTTCATAGCTCAAGCTGTAGATCTCAAGTTCCGATGATGATGATGATGATGATGATG 1503
Qy 2161 GGGCTGTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATTCGATGATGATGATG 2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATTCGATGATGATGATG 1443
Qy 2221 TCTTGAATCAACCGGTAAGACAGCACTTATGCGCACTGCAAGACGCACTGATGACAG 2280
Db 1442 TCTTGAATCAACCGGTAAGACAGCACTTATGCGCACTGCAAGACGCACTGATGACAG 1383

QY 2281 GATTAGCAGAGCGAGGATGTAGCGGCTGACAGAGTCTTGAAGTGGTGGCTTAAGTACTA 2340
 DB 1382 GATTAGCAGAGCGAGGATGTAGCGGCTGACAGAGTCTTGAAGTGGTGGCTTAAGTACTA 1323
 QY 2241 CGGCTACACTAGAGGACAGTATTTGTATCTGCGCTGCTGCTGAAGCCAGTTACCTTCGG 2400
 DB 1322 CGGCTACACTAGAGGACAGTATTTGTATCTGCGCTGCTGCTGAAGCCAGTTACCTTCGG 1263
 QY 2401 AAAAAGATGTAGAGTCTTGAATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTT 2460
 DB 1262 AAAAAGATGTAGAGTCTTGAATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTT 1203
 QY 2461 TGTTCGACAGCAGAGATTTACGCGAGAAAAGAGATCTCAAGATCTTTGATCTT 2520
 DB 1202 TGTTCGACAGCAGAGATTTACGCGAGAAAAGAGATCTCAAGATCTTTGATCTT 1143
 QY 2521 TTTCAAGGGGCTGACGCTCAGTGGACGAAAACCTACGTTAAGGGATTTTGGTCATGAG 2580
 DB 1142 TTTCAAGGGGCTGACGCTCAGTGGACGAAAACCTACGTTAAGGGATTTTGGTCATGAG 1083
 QY 2581 ATTATCAAAAAGATCTTCACTAGATCTTTTAATTAATTAAGATTTTAATCAAT 2640
 DB 1082 ATTATCAAAAAGATCTTCACTAGATCTTTTAATTAATTAAGATTTTAATCAAT 1023
 QY 2641 CTTAAGATATATAGATTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCAC 2700
 DB 1022 CTTAAGATATATAGATTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCAC 963
 QY 2701 TATCTCAGCGATCTGTCTATTTGCTTATCATCATAGTTGCTGACCTCCGCTGCTAGAT 2760
 DB 962 TATCTCAGCGATCTGTCTATTTGCTTATCATCATAGTTGCTGACCTCCGCTGCTAGAT 903
 QY 2761 AACTACGATACGGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACC 2820
 DB 902 AACTACGATACGGGAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACC 843
 QY 2821 ACCGCTACCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGGCCGAG 2880
 DB 842 ACCGCTACCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGGCCGAG 783
 QY 2881 AAGTGTCTCGCAACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAGGTAG 2940
 DB 782 AAGTGTCTCGCAACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAGGTAG 723
 QY 2941 AGTAAGTGTGCGCAGTTAATAGTTTSCGCAAGTGTGGCCATCTCTACAGCATCT 3000
 DB 722 AGTAAGTGTGCGCAGTTAATAGTTTSCGCAAGTGTGGCCATCTCTACAGCATCT 663
 QY 3001 GGTGTACGCTGCTGCTTGTGTAGGCTTCACTTCACTCCGCTTCCCAAGCATCAAGGG 3060
 DB 662 GGTGTACGCTGCTGCTTGTGTAGGCTTCACTTCACTCCGCTTCCCAAGCATCAAGGG 603
 QY 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGGGTTAGCTCTTGGTCTCCGATCT 3120
 DB 602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGGGTTAGCTCTTGGTCTCCGATCT 543
 QY 3121 TGTGAGAGTATGTGCGCCAGTGTATATCATCTATGTTATGCGCAGCTCATATATTC 3180
 DB 542 TGTGAGAGTATGTGCGCCAGTGTATATCATCTATGTTATGCGCAGCTCATATATTC 483
 QY 3181 TCTTACTGTATGCGCATCCGTAAGATGCTTTCTGTGAGTGTGTAGTCTCAACCAATC 3240
 DB 482 TCTTACTGTATGCGCATCCGTAAGATGCTTTCTGTGAGTGTGTAGTCTCAACCAATC 423
 QY 3241 ATTCTGAGATATGTATGTGCGCAGAGTGTCTTGGCCGCGGTCAATATAGGAGTAA 3300
 DB 422 ATTCTGAGATATGTATGTGCGCAGAGTGTCTTGGCCGCGGTCAATATAGGAGTAA 363
 QY 3301 TACCGGCGCATATAGCAGAACTTTAAAGTGTCTCATCTGAGAAAAGTTTCTGCGGCG 3360
 DB 362 TACCGGCGCATATAGCAGAACTTTAAAGTGTCTCATCTGAGAAAAGTTTCTGCGGCG 303

QY 3361 AAAACTCTCAAGATCTTACCGCTGTGAATTCAGTTCAATGTAACCCACTGTCACC 3420
 DB 302 AAAACTCTCAAGATCTTACCGCTGTGAATTCAGTTCAATGTAACCCACTGTCACC 243
 QY 3421 CACTGATCTTACAGATCTTTTACTTTCACACAGGTTCTTGAGTGAGCAAAAACAGAG 3480
 DB 242 CACTGATCTTACAGATCTTTTACTTTCACACAGGTTCTTGAGTGAGCAAAAACAGAG 183
 QY 3481 GCAAAATGCGCAAAAAGGAAATTAAGGGGAGACACGAAATGTTGAATCTCATCTCTT 3540
 DB 182 GCAAAATGCGCAAAAAGGAAATTAAGGGGAGACACGAAATGTTGAATCTCATCTCTT 123
 QY 3541 CCTTTTCAATATTTATGAAGATTTATCAGGGTTATGTCTCATGAGCGGATACATATT 3600
 DB 122 CCTTTTCAATATTTATGAAGATTTATCAGGGTTATGTCTCATGAGCGGATACATATT 63
 QY 3601 TGAATGATTTAGAAAAATTAACCAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCC 3660
 DB 62 TGAATGATTTAGAAAAATTAACCAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCC 3
 QY 3661 AC 3662
 DB 2 AC 1

RESULT 14
 US-10-637-758-17/c
 / Sequence 17, Application US/10637758
 / Publication No. US20040180352A1
 / GENERAL INFORMATION:
 / APPLICANT: Large Scale Biology Corporation
 / TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
 / FILE REFERENCE: LSBC-PIG5392-CIP
 / CURRENT APPLICATION NUMBER: US/10/637,758
 / PRIOR FILING DATE: 2003-08-08
 / PRIOR APPLICATION NUMBER: US/10/280,913
 / PRIOR FILING DATE: 2002-10-25
 / PRIOR APPLICATION NUMBER: 60/402,342
 / PRIOR FILING DATE: 2002-08-08
 / PRIOR APPLICATION NUMBER: 10/066,390
 / PRIOR FILING DATE: 2002-02-01
 / PRIOR APPLICATION NUMBER: 60/268,785
 / PRIOR FILING DATE: 2001-02-14
 / PRIOR APPLICATION NUMBER: 60/266,386
 / PRIOR FILING DATE: 2002-02-02
 / NUMBER OF SEQ ID NOS: 35
 / SOFTWARE: Patentin version 3.1
 / SEQ ID NO 17
 / LENGTH: 3637
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Encodes cycle 3 GFP
 US-10-637-758-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;
 Best Local Similarity 83.8%; Pred. No. 0;
 Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

QY 1 CTTAATTTGTAAGCGTTAATATTTTGTAAATTCGCTTAATTTTGTAAATCGCTC 60
 DB 3637 CTTAATTTGTAAGCGTTAATATTTTGTAAATTCGCTTAATTTTGTAAATCGCTC 3578
 QY 61 ATTTTAAACCAATAGGCGGAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGACGA 120
 DB 3577 ATTTTAAACCAATAGGCGGAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGACGA 3518
 QY 121 GATAGGGTGAAGTGTGTTCAGTTTGAACAAGAGTCCACTATTAAAGAGTGAATC 180
 DB 3517 GATAGGGTGAAGTGTGTTCAGTTTGAACAAGAGTCCACTATTAAAGAGTGAATC 3458
 QY 181 CAACTCAAGAGGCGGAAAAAGCGTATACAGGCGATGCGCCACTAGTGAACATCAACC 240

Db 3457 CAACGTCAAGGCGGAAAAACGCTATCAGGGCGATGGCCCACTACGTAACATCAC 3398
Qy 241 CTAATCAAGTTTTTGGGGTGAAGTCCGTAAGCACTAAATCGGAACCTTAAAGGAG 300
Db 3397 CTAATCAAGTTTTTGGGGTGAAGTCCGTAAGCACTAAATCGGAACCTTAAAGGAG 3338
Qy 301 CCCCCGATTTAGAGCTTGAACGGGGGAAAGCGGCGAAACGTGGCGAAGAAAGGAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGAACGGGGGAAAGCGGCGAAACGTGGCGAAGAAAGGAGAA 3278
Qy 361 AGCGAAGAGAGCGGGCGCTGAGGCGCTGCAAGTACGCTGCAAGTCCGCTAAACAC 420
Db 3277 AGCGAAGAGAGCGGGCGCTGAGGCGCTGCAAGTACGCTGCAAGTCCGCTAAACAC 3218
Qy 421 CACACCCCGCGCTTAATGCGCCGCTACAGGGCCGCTCCCATTTGCCATTCAGCTCG 480
Db 3217 CACACCCCGCGCTTAATGCGCCGCTACAGGGCCGCTCCCATTTGCCATTCAGCTCG 3158
Qy 481 CAACCTGTTGGGAAGGGCGATCGGTCGGGCGCTTCGCTAATTAAGCCAGCTGGCGAAG 540
Db 3157 CAACCTGTTGGGAAGGGCGATCGGTCGGGCGCTTCGCTAATTAAGCCAGCTGGCGAAG 3098
Qy 541 GGGATGCTGCTCAAGCGCATTAAGTTGGGTAACGCGCAGGGTTTTCCAGTCAAGCTTG 600
Db 3097 GGGATGCTGCTCAAGCGCATTAAGTTGGGTAACGCGCAGGGTTTTCCAGTCAAGCTTG 3038
Qy 601 TAAAAAGACGGCCAGTGAAGCGCGCTCGTTCACTTACGTTTTGAACCCGTGGAGAGCG 660
Db 3037 TAAAAAGACGGCCAGTGAAGCGCGCTCGTTCACTTACGTTTTGAACCCGTGGAGCTCA 2978
Qy 661 GCAGACTCGGAGTGCMAATGTGTTTACAGCGTATGAGCAGATGAAGATGCTGAGAC 720
Db 2977 CCGGCGTGGCGG-----CGGCTTGAAGCTAAGTGAATCCG--CGG 2939
Qy 721 GCTGCGAAGACGCGAGCTAGATTAACCTAGAAAGATATCATATTGTAAGTAAGTAA 780
Db 2938 GCTGCGAAGATTTCTTATTTGTAATAGTTCATCCATCCATGTAATCCAGAGAGATTA 2879
Qy 781 AGATTAATATGCTGTAATTTGAACGATGGGATCTGTAAATGACTCACTAATAGGCGGAT 840
Db 2878 CAACCTCAAGAGGACCAATGTGCTCACGCTTTTCGTTGGGATCTTTCCAAAGGCGAGATT 2819
Qy 841 TGGGATCGGGGCGGCGGCTCGAGGTGACGAGTATGATAGTTGATATGAAATTCCTGC 900
Db 2818 GTTGCGAAGGTAATGTTGTCTGTGTAAGAGACAGGGCCATCGCAATTTGAGATTTT 2759
Qy 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACGCGGTGAGCTCAGCTTTTGT 960
Db 2758 GTTGATATGATGCTGCTAGTTGAACGATCCATCTTCAATGTGTGGGAAATTTGAAGT 2699
Qy 961 TCCCTTAATGAGGGTTAATTAATCCATGCTCAATTTTACGCAACTAATCTTTCTAG 1020
Db 2698 TAGCTTGAATTCATCTTTGTTGTGTGCGGTATGATATCAATTTGTAG--TTATAG 2640
Qy 1021 GATTATTAAGTATGATAGATATATGATATGATGAGTCTTTTCCGCTCAGTATCCG 1080
Db 2639 TTGATCTGAAGTTGTGTGCGAAGATTTTCATCTTTTAAATCAATACCTTTTAAC 2580
Qy 1081 CAAGCTGCGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCGCGAGGTTG 1140
Db 2579 TCGATAGATTAACAGGGGTAACCTTCAAACTTGAATTCAGACGCGGTCTTGTAGTTC 2520
Qy 1141 AAGCGGATGAGAAAGATTTGCCGAGATGACTGCTGCTGATGAAGCTTGAAGGAAAC 1200
Db 2519 CCGTCATCTTTGAAGATATGATGTGCTCTGTATCAATTAACCTTGGGATGACCTCTTG 2460
Qy 1201 GCAGCTTACCATGATGATTCGGGAAGGTGCGCATCAGCCCTTAAAGGTGAAGT 1260
Db 2459 AAAAAGTCATCCGTTTATATGATCCGATTAACGGGAAGACATTTGAACCATTAAGAG 2400
Qy 1261 TCGTTACAGGCACTGGGATACAGATTGCTGCGGCTTTTCCGACACAGTTCCGAGTGG 1320
Db 2399 AAAGTAGACAAAGTGTGGCGATGAACAGGTAGTTTTCCAGTAGTGCATAATAATTTA 2340

Qy 1321 TCAGCCGAAAGCGATACGCAACCCGAACATATCCGCGAGACGCCGGAATCGCGTGGC 1380
Db 2339 AGGTAAGCTTTCGTAATGTAGCATCACTTACCCCTCTCACTGACAGAAATTTGTGC 2280
Qy 1381 GTGTGAGATTAATGACAGCGGTGCGGCGCTGGGATATTAAGTACAGGAGACGGGTATC 1440
Db 2279 CCATTAACATCAATCTAATTTCAACAGAAATTTGGACAACTCCAGTAAAAATTTCTT 2220
Qy 1441 CTGCGTGAATCCGAGAAATGACATGATATCCCGTGAATTAACCGCGCGGCGCTT 1500
Db 2219 C---CTTACTCATCGGTAACCAAGCTTTTGTCCCTTATGAGAGGTTAATTTGCGCGCTT 2163
Qy 1501 GCGCTAATCATGCTCATGAGCTTTTCTGTGTGAATTTGTAATCGCTCAATTTCCCA 1560
Db 2162 GCGCTAATCATGCTCATGAGCTTTTCTGTGTGAATTTGTAATCGCTCAATTTCCCA 2103
Qy 1561 CAACATACGACCGGAAAGCATPAAAGTGTAAAGCTGGGGTCTTAATGATGAGTAACT 1620
Db 2102 CAACATACGACCGGAAAGCATPAAAGTGTAAAGCTGGGGTCTTAATGATGAGTAACT 2043
Qy 1621 CACATTAATTCGTTGCGCTCACTGCGCTTTTCAAGTCGAGAAACCTGTGTCAGACT 1680
Db 2042 CACATTAATTCGTTGCGCTCACTGCGCTTTTCAAGTCGAGAAACCTGTGTCAGACT 1983
Qy 1681 GCATTAATGATTCGCGCAACGCGGGGAGAGCGGCTTTGCTATTTGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGATTCGCGCAACGCGGGGAGAGCGGCTTTGCTATTTGGCGCTCTTCCGC 1923
Qy 1741 TTCTCGCTCACTGACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1922 TTCTCGCTCACTGACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
Qy 1801 CTCAAGCGGTAATACGTTATTCACAGAAATCAGGGGATTAACGAGAAAGAAATGTG 1860
Db 1862 CTCAAGCGGTAATACGTTATTCACAGAAATCAGGGGATTAACGAGAAAGAAATGTG 1803
Qy 1861 AGCAAAAGCGGCAAGAAAGGCGAGAAACCGTAAAGAGCGCGCTTTGCTGCGTTTTTCA 1920
Db 1802 AGCAAAAGCGGCAAGAAAGGCGAGAAACCGTAAAGAGCGCGCTTTGCTGCGTTTTTCA 1743
Qy 1921 TAGGCTCGGCGGCTGAGCGAGTCAAAAAATCGACGCTCAAGTCAAGAGTGGCGAAA 1980
Db 1742 TAGGCTCGGCGGCTGAGCGAGTCAAAAAATCGACGCTCAAGTCAAGAGTGGCGAAA 1683
Qy 1981 CCGGACAGGACTAATAAGATACAGAGGCTTCCCTGGAAGCTCCCTGTCGCTCTCC 2040
Db 1682 CCGGACAGGACTAATAAGATACAGAGGCTTCCCTGGAAGCTCCCTGTCGCTCTCC 1623
Qy 2041 TGTTCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAAGCTGGC 2100
Db 1622 TGTTCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAAGCTGGC 1563
Qy 2101 GCTTCTCATAGCTCAGGCTGATGATTCAGTTCGATGAGTTCGTTGCTGCTCAAGCT 2160
Db 1562 GCTTCTCATAGCTCAGGCTGATGATTCAGTTCGATGAGTTCGTTGCTGCTCAAGCT 1503
Qy 2161 GGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGGTCGCGCTTATCCGGAATCTATCG 2220
Db 1502 GGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGGTCGCGCTTATCCGGAATCTATCG 1443
Qy 2221 TCTTGAATCCACCCGGAAGACAGACTTAATGCGCACTGGCAGACGACTGTAACAG 2280
Db 1442 TCTTGAATCCACCCGGAAGACAGACTTAATGCGCACTGGCAGACGACTGTAACAG 1383
Qy 2281 GATTAGCAGACGAGTATGATGAGCGGTGCTACAGATTTCTTGAAGTGTGGCTTACTA 2340
Db 1382 GATTAGCAGACGAGTATGATGAGCGGTGCTACAGATTTCTTGAAGTGTGGCTTACTA 1323
Qy 2341 CCGCTACACTAAGAGCAGTATTTGATCTGCGCTCTGCTGAAGCAGTAACTTCTCG 2400
Db 1322 CCGCTACACTAAGAGCAGTATTTGATCTGCGCTCTGCTGAAGCAGTAACTTCTCG 1263

QY 2401 AAAAGAGTTGAGTCTCTTGAATCCGGCAAAACCAACACCGCTGTAGCGGTGTTTTT 2460
 Db 1262 AAAAGAGTTGAGTCTCTTGAATCCGGCAAAACCAACACCGCTGTAGCGGTGTTTTT 1203
 QY 2461 TGTTCGACGACGAGATTAAGCGCAGAAAAAGATCTCAAGAAATCTTTGATCTT 2520
 Db 1202 TGTTCGACGACGAGATTAAGCGCAGAAAAAGATCTCAAGAAATCTTTGATCTT 1143
 QY 2521 TTTACGGGGTCTGACGCTCAGTGAAGCAAAAATCAGTTAAGGATTTTGGTATGAG 2580
 Db 1142 TTTACGGGGTCTGACGCTCAGTGAAGCAAAAATCAGTTAAGGATTTTGGTATGAG 1083
 QY 2581 ATTATCAAAAAGATCTTCACTAGATCTTTTAATTAATAATGAATTTAATCAAT 2640
 Db 1082 ATTATCAAAAAGATCTTCACTAGATCTTTTAATTAATAATGAATTTAATCAAT 1023
 QY 2641 CTAAAGTATATATAGTAAACTTGTCTCAAGTTACCAATGCTTAATCAGTGAAGCACC 2700
 Db 1022 CTAAAGTATATATAGTAAACTTGTCTCAAGTTACCAATGCTTAATCAGTGAAGCACC 963
 QY 2701 TATCTAGGATCTGTCTATTTTCGTTCAATCCATAGTTGCCGATCTCCCGCTGTAGAT 2760
 Db 962 TATCTAGGATCTGTCTATTTTCGTTCAATCCATAGTTGCCGATCTCCCGCTGTAGAT 903
 QY 2761 AACTACGATACGGAGGGCTTACCAATCGGCCCAAGTGTGCAATGATACCGGAGACC 2820
 Db 902 AACTACGATACGGAGGGCTTACCAATCGGCCCAAGTGTGCAATGATACCGGAGACC 843
 QY 2821 ACGCTACCGGCTCAATTTATGACAAATTAACAGCCAGCCGGAAGGGCCGACGCGAG 2880
 Db 842 ACGCTACCGGCTCAATTTATGACAAATTAACAGCCAGCCGGAAGGGCCGACGCGAG 783
 QY 2881 AAGTGTCTGCACTTATTCGCTCCATCCAGTCTATTAATGTTGCGGGAAGCTAG 2940
 Db 782 AAGTGTCTGCACTTATTCGCTCCATCCAGTCTATTAATGTTGCGGGAAGCTAG 723
 QY 2941 AGTAAGTATGTCGCACTTATGTTGCGCAACGTTGCGCAATGCTAAGCAATCT 3000
 Db 722 AGTAAGTATGTCGCACTTATGTTGCGCAACGTTGCGCAATGCTAAGCAATCT 663
 QY 3001 GGTGTCAAGCTGCTGTTGATGATGAGCTTCAATGAGCTCCGTTCCCAAGATCAAGCG 3060
 Db 662 GGTGTCAAGCTGCTGTTGATGATGAGCTTCAATGAGCTCCGTTCCCAAGATCAAGCG 603
 QY 3061 AGTTACATGATCCCAATGTTGCAAAAAAGCGTTAGCTCTTGCGTCTCCGATCT 3120
 Db 602 AGTTACATGATCCCAATGTTGCAAAAAAGCGTTAGCTCTTGCGTCTCCGATCT 543
 QY 3121 TGTCAAGATTAAGTGGCGCAGTGTATCATCATAGTGTATGCGCATATATTC 3180
 Db 542 TGTCAAGATTAAGTGGCGCAGTGTATCATCATAGTGTATGCGCATATATTC 483
 QY 3181 TCTTACTGTCAATGCGCATCCGTAAGATGCTTTGCTGTGACTGTGTAGTATCAACCAATGC 3240
 Db 482 TCTTACTGTCAATGCGCATCCGTAAGATGCTTTGCTGTGACTGTGTAGTATCAACCAATGC 423
 QY 3241 ATTCTGAGATTAAGTGTGCGGCAACGAGTGTCTTGCCCGCGGTCAATACGGAATTA 3300
 Db 422 ATTCTGAGATTAAGTGTGCGGCAACGAGTGTCTTGCCCGCGGTCAATACGGAATTA 363
 QY 3301 TACCGCGCACAATAGCAAGACTTTAAAGTCTCATCATGTTGGAAGAGTCTTCCGAGGAG 3360
 Db 362 TACCGCGCACAATAGCAAGACTTTAAAGTCTCATCATGTTGGAAGAGTCTTCCGAGGAG 303
 QY 3361 AAAAATCTCAAGATTTTACCGCTGTGAGATCCAGTTGCAATGTAACCCACTGTGACAC 3420
 Db 302 AAAAATCTCAAGATTTTACCGCTGTGAGATCCAGTTGCAATGTAACCCACTGTGACAC 243
 QY 3421 CAATGATCTTCAAGATCTTTTACCGCTGTGAGATCCAGTTGCAATGTAACCCACTGTGACAC 3480
 Db 242 CAATGATCTTCAAGATCTTTTACCGCTGTGAGATCCAGTTGCAATGTAACCCACTGTGACAC 183
 QY 3481 GCAAAATCCGCAAAAAAGGAATTAAGGCGCAACGGAATGTTGAATCTACTCTT 3540

Db 182 GCAAAATGCGCAAAAAAGGAATTAAGGCGCAACGGAATGTTGAATCTACTCTT 123
 QY 3541 CCTTTTCAATTTTATTAAGCACTTTTATGAGTTATGTTGATGAGGAGTATTAATTT 3600
 Db 122 CCTTTTCAATTTTATTAAGCACTTTTATGAGTTATGTTGATGAGGAGTATTAATTT 63
 QY 3601 TGAATGTTTATTAAGAAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGAAAGTCC 3660
 Db 62 TGAATGTTTATTAAGAAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGAAAGTCC 3
 QY 3661 AC 3662
 Db 2 AC 1

RESULT 15
 US-10-066-390-3/c
 ; Sequence 3, Application US/10066390
 ; Publication No. US20020146732A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padgett, Hal S.
 ; APPLICANT: Lindbo, John A.
 ; TITLE OF INVENTION: A Method of Increasing Complementarity
 ; FILE REFERENCE: P-UG 4878
 ; CURRENT APPLICATION NUMBER: US/10/066,390
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 3637
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 US-10-066-390-3

Query Match 72.5%; Score 2656.4; DB 13; Length 3637;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;
 QY 1 CTAAATGTAACGCTTAATTTGTTAAATGCGCTTAATTTGTTAAATGAGCTC 60
 Db 3637 CTAAATGTAACGCTTAATTTGTTAAATGCGCTTAATTTGTTAAATGAGCTC 3578
 QY 61 ATTTTAAACCAATAGGCGCAAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGCCGA 120
 Db 3577 ATTTTAAACCAATAGGCGCAAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGCCGA 3518
 QY 121 GATAGGTTGAGTGTTCCTGATTTGGAACAAGAGTCCACTTAATTAAGAAAGTGAATC 180
 Db 3517 GATAGGTTGAGTGTTCCTGATTTGGAACAAGAGTCCACTTAATTAAGAAAGTGAATC 3458
 QY 181 CAAGTCAAAAGGGGCAAAAACGCTCATAGGGCGATGGCCCACTAAGTGAACCAATCACC 240
 Db 3457 CAAGTCAAAAGGGGCAAAAACGCTCATAGGGCGATGGCCCACTAAGTGAACCAATCACC 3398
 QY 241 CTATCAAGTTTGTGGGTCGAGGTGCGTAAACATCAATATGGAACCTTAAGGAG 300
 Db 3397 CTATCAAGTTTGTGGGTCGAGGTGCGTAAACATCAATATGGAACCTTAAGGAG 3338
 QY 301 CCCCCGATTAGAGCTTGAACGGGAAACCGCGCAACGTGCGCAAGAAAGGAAGGAAGAA 360
 Db 3337 CCCCCGATTAGAGCTTGAACGGGAAACCGCGCAACGTGCGCAAGAAAGGAAGGAAGAA 3278
 QY 361 AGCGAAAGAGGCGCGCTTGAAGGCGCTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
 Db 3277 AGCGAAAGAGGCGCGCTTGAAGGCGCTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3218
 QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCCATTCGCAATGAGGCTGCG 480

Dh 3217 CACACCGCGCGGCTTAATGCGCGCTACAGGGCGCGTCCATTGCGCAATTGAGGCTGCG 3158
Qy 481 CAACCTTTGGGAAGGCGCATGCGTGGCGGCTCTTGCTATTAGCGCACTGGCGAAGG 540
Dh 3157 CAACCTTTGGGAAGGCGCATGCGTGGCGGCTCTTGCTATTAGCGCACTGGCGAAGG 3098
Qy 541 GGGATGTCGTCAGAGGCGATTAGTTGGTTAAACGCAAGGTTTCCAGTCAGCAAGCTTG 600
Dh 3097 GGGATGTCGTCAGAGGCGATTAGTTGGTTAAACGCAAGGTTTCCAGTCAGCAAGCTTG 3038
Qy 601 TAAACGACGCGCAGTGAAGCGCGCTCGTTCAATTCACGTTTGTAAACCGCGAGAGCG 660
Dh 3037 TAAACGACGCGCAGTGAAGCGCGGCTATACGACTCACTATAGGCGCAATTGGAGCTCA 2978
Qy 661 GCAAGCTGCGCGTGCMAATGTGTTTACAGCGTGAAGAGATGAAGATGCTGACAC 720
Dh 2977 CCGCGGTGGCGCGCTGTAACACTAGTGAATCCCGCGGCTGACGAAATCTTAATTTGT 2918
Qy 721 GCTGACAGACAGCAGCTAGATTAACTTGAAGATATATATGTAAGTACGTTAA 780
Dh 2917 ATAGTTCAATCCATGCGCATGTGTAAATCCAGCAGCTGTAAACAACCTCAAGAGACATGT 2858
Qy 781 AGATATATC--ATGCGTAAATTTGAAGCATGGGATCTGTAAATACGACTCATATAGCGCA 838
Dh 2857 GGTCTCTTTTCTGTTGGGATCTTTGAAAGGCGAGATGTGTGAGACAGTAATGCTGT 2798
Qy 839 ATTGGGTAACGGGCGCGCGCTCGAGGTCGACGTAATGTAAGCTTGATTCGAAATTCCT 898
Dh 2797 CTGCTAAAGACAGAGGCGCATGCGCAATTGGATATTTGTTGATTAATGCTGCTAGTT 2738
Qy 899 GCAGCGCGGGGATTCATAGTTCTAGAGCGCGCGCACCGCGGTGAGCTCGAGCTTT 958
Dh 2737 GAAAGCTTCAATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCATCTTTT 2678
Qy 959 GTTCCCTTTAGTGAAGGTTAATTAATGATCCATGCGCTAATTTTACGACGATCTTCT 1018
Dh 2677 GTTGTCTGCATGATGATATATGTTGTGA--GTTATAGTTGTATTCGAATTTGTGTCA 2619
Qy 1019 AGGTTAATCTAGCTGATCAGATCATATGTCGGGCTTTTTCGCGCTCAGTCATCG 1078
Dh 2618 AGATGTTTCATCTTCTTAAATCAATACCTTTTAACTCGATTC-----TATTAACA 2564
Qy 1079 CCCAAGCTGGGCTATCTGGGCTATCGGGCATCGGGAGGAAGAACCGCTGCTTTTCCGCGAGT 1138
Dh 2563 GGGTATACCTTCAAACTTGACTTC-----AGCAGTGTCTTGTAGTTCCCTTC 2515
Qy 1139 TGAAGCGCATGAAAGAGTTTGCAGAGATGACTGCTGATTAAGCTTGAAGCGAA 1198
Dh 2514 ATCTTGAATAATATATGTTCTTCTGTACATTAACCTTCGGGATGCACTCTTGAATA 2455
Qy 1199 ACGCAGCTTAACTATGATATTCGGGAAGTGTGGCCATGACCGCTTAAACGCTGAAT 1258
Dh 2454 GTCATGCGGTTTCAATGATCTGGGATCTTGAAGAACATGAAACACATTAAGAAAGT 2395
Qy 1259 GTTGTTCAGGCGCATCGGGATACAGTTGCTGGGCTTTTCCGGAACAGTTCCGAT 1318
Dh 2394 AGTACAAGTGTGGCCATGGAACAGTATTTTCCAGTATGCAATAATTAATTTAAGGT 2335
Qy 1319 GGTGAGCCGGAAGGCGCATCAGCAACCGGAACAATACCGGCGACAGCGGAACTGCGGTC 1378
Dh 2334 AAGTTTTC-----GTAATGTTCATCACTTCACTTCCATGACGAAATTTGTGTC 2280
Qy 1379 CGGTGTGCAATTAATGACAGCGGTGCGGCTGGGATATTACGTACAGCAGAGAGGTA 1438
Dh 2279 CCATTAACTACATCTAATTAATTAACAAGAAATTTGGGACAACTCCAGTGAATAAGTTCTTCT 2220
Qy 1439 TCCGTGCGTGAATGCGGAGAAATGGAATGATACCCCGTGAATTAACCGCGCGCGCGC 1498
Dh 2219 CTTTACTCATCGGTACCACTTTTGTCTCTTATAGTGAAGGTTA-----ATTGCGGC 2165
Qy 1499 TTGCGTGAATCATGCTGATGCTGTTCTGTGTGAATTTGTAATCCGCTCACAATTTCA 1558
Dh 2164 TTGGGTAATCATGCTGATGCTGTTCTGTGTGAATTTGTAATCCGCTCACAATTTCA 2105

Qy 1559 CACAACATACGAGCGCGGAAGATTAAGTTAAACCTGGGGTGCCTAATGATGAGCTAA 1618
Dh 2104 CACAACATACGAGCGCGGAAGATTAAGTTAAACCTGGGGTGCCTAATGATGAGCTAA 2045
Qy 1619 CTCACTTAATTTGGGTTGCGCTCACTGCGCGCTTCCAGTGGGAAACCTGTCGTCAG 1678
Dh 2044 CTCACTTAATTTGGGTTGCGCTCACTGCGCGCTTCCAGTGGGAAACCTGTCGTCAG 1985
Qy 1679 CTGCAATTAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTGCGTATTGGGCGCTTTC 1738
Dh 1984 CTGCAATTAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTTGCCTATTGGGCGCTTTC 1925
Qy 1739 GCTTCTGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCGGCGGATCACT 1798
Dh 1924 GCTTCTGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCGGCGGATCACT 1865
Qy 1799 CACTCAAAAGCGGTAATACGGTTATCCAGATACAGAGGATTAACGAGAAAGACATG 1858
Dh 1864 CACTCAAAAGCGGTAATACGGTTATCCAGATACAGAGGATTAACGAGAAAGACATG 1805
Qy 1859 TGAACAAAGGCGCAGAAAGGCGGAAACCGTAAAGAGCGGTTGCTGCGCTTTTC 1918
Dh 1804 TGAACAAAGGCGCAGAAAGGCGGAAACCGTAAAGAGCGGTTGCTGCGCTTTTC 1745
Qy 1919 CATAGGCTCGCGCGCGCTGACAGATCAAAAAATGACGCTCAAGTCAAGTGCAG 1978
Dh 1744 CATAGGCTCGCGCGCGCTGACAGATCAAAAAATGACGCTCAAGTCAAGTGCAG 1685
Qy 1979 AACCGGACAGGACTATTAAGATACAGAGCGTTTCCCGTGAACCTCCTGTCGCT 2038
Dh 1684 AACCGGACAGGACTATTAAGATACAGAGCGTTTCCCGTGAACCTCCTGTCGCT 1625
Qy 2039 CCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCCGGAAGCGTG 2098
Dh 1524 CCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCCGGAAGCGTG 1565
Qy 2099 GCGCTTCTCATACCTCAACGCTGATGATCTCACTGTTGCTGTTGCTGCTCAAG 2158
Dh 1564 GCGCTTCTCATACCTCAACGCTGATGATCTCACTGTTGCTGTTGCTGCTCAAG 1505
Qy 2159 CTGGGCTGTGTCAGAAACCCCGCTCAGCGCGACCGCTGCGCTTATCCGGTAATAT 2218
Dh 1504 CTGGGCTGTGTCAGAAACCCCGCTCAGCGCGACCGCTGCGCTTATCCGGTAATAT 1445
Qy 2219 CGTCTGATGTCACACCGGTAAGACAGACTTATCGCACTGAGCAGCAGCACTGTATAC 2278
Dh 1444 CGTCTGATGTCACACCGGTAAGACAGACTTATCGCACTGAGCAGCAGCACTGTATAC 1385
Qy 2279 AGATTTAGCAGAGCGAGATATGAGCGGTGCTACAGAGTTCTTGAAGTGTGCTTAC 2338
Dh 1384 AGATTTAGCAGAGCGAGATATGAGCGGTGCTACAGAGTTCTTGAAGTGTGCTTAC 1325
Qy 2339 TAGGCTTACCTAAGACAGATTTGTTATCTGCGCTCTGCGAAGCGCATTAACCTTC 2398
Dh 1324 TAGGCTTACCTAAGACAGATTTGTTATCTGCGCTCTGCGAAGCGCATTAACCTTC 1265
Qy 2399 GGAAGAAAGATGTTGTTACTCTTGTATCCGGCAACAAACACCGCTGATGAGGTTT 2458
Dh 1264 GGAAGAAAGATGTTGTTACTCTTGTATCCGGCAACAAACACCGCTGATGAGGTTT 1205
Qy 2459 TTTGTTTGAAGCAGCAGATTAACGCGGAGAAAGAAAGATCTCAAGAGATCTTTATC 2518
Dh 1204 TTTGTTTGAAGCAGCAGATTAACGCGGAGAAAGAAAGATCTCAAGAGATCTTTATC 1145
Qy 2519 TTTTCTAAGGCTGTGACGCTCAGTGAAGAAATCAGCTTAAGGATTTTGGTCATG 2578
Dh 1144 TTTTCTAAGGCTGTGACGCTCAGTGAAGAAATCAGCTTAAGGATTTTGGTCATG 1085
Qy 2579 AGATTATCAAAAGAGATCTTCACTAGATCTTTTAATTAATGAATGATTTTAAATCA 2638
Dh 1084 AGATTATCAAAAGAGATCTTCACTAGATCTTTTAATTAATGAATGATTTTAAATCA 1025

QY 2639 ATCTAAGTATATATAGTAAACTTGGTCTGACAGTTTACATCTTAAATCAGTAGCA 2698
| | | | |
Db 1024 ATCTAAGTATATATAGTAACTTGGTCTGACAGTTTACATCTTAAATCAGTAGCA 965
| | | | |
QY 2699 CCTATCTCAGCGATCTGTCTATTTTCGTCATCTCAATAGTGGCTGACCTCCCGTCGTGTAG 2758
| | | | |
Db 964 CCTATCTCAGCGATCTGTCTATTTTCGTCATCTCAATAGTGGCTGACCTCCCGTCGTGTAG 905
| | | | |
QY 2759 ATTAAGTATACGGAGAGGGCTTACATCTGAGCCCGAGTCGTCAATGATACCGCAGAC 2818
| | | | |
Db 904 ATTAAGTATACGGAGAGGGCTTACATCTGAGCCCGAGTCGTCAATGATACCGCAGAC 845
| | | | |
QY 2819 CCACGCTCACCGGCTCCAGATTATACAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCCG 2878
| | | | |
Db 844 CCACGCTCACCGGCTCCAGATTATACAGCAATAAACCCAGCAGCCGGAAGGGCCGAGCCG 785
| | | | |
QY 2879 AGAAGTGTCTCTGCACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCT 2938
| | | | |
Db 784 AGAAGTGTCTCTGCACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCT 725
| | | | |
QY 2939 AGAAGTATAGTTCCGCAATTAAATAGTTTCCGCAAGTTGTGCCATTGCTACAGCATC 2998
| | | | |
Db 724 AGAAGTATAGTTCCGCAATTAAATAGTTTCCGCAAGTTGTGCCATTGCTACAGCATC 665
| | | | |
QY 2999 GTGGTGCAGCGTCGTGCTTGGTATGGCTTCAATTCAGCTCCGCTTCCCAAGATCAAG 3058
| | | | |
Db 664 GTGGTGCAGCGTCGTGCTTGGTATGGCTTCAATTCAGCTCCGCTTCCCAAGATCAAG 605
| | | | |
QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTGGTCTCCGATC 3118
| | | | |
Db 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTGGTCTCCGATC 545
| | | | |
QY 3119 GTTGTCAAGTAAAGTGGCCCGCAGTGTATCATCTCATGTTATGGCAGACCTGCATAT 3178
| | | | |
Db 544 GTTGTCAAGTAAAGTGGCCCGCAGTGTATCATCTCATGTTATGGCAGACCTGCATAT 485
| | | | |
QY 3179 TCTTACTGTGATGCAATCCGTAGATGCTTTCTGTGACTGTGTGAGTACTCAACCAAG 3238
| | | | |
Db 484 TCTTACTGTGATGCAATCCGTAGATGCTTTCTGTGACTGTGTGAGTACTCAACCAAG 425
| | | | |
QY 3239 TCAATCTGAGAAATGATGTGGGCGAGCCAGTGTCTTGGCCGCGCTCAATACCGGAT 3298
| | | | |
Db 424 TCAATCTGAGAAATGATGTGGGCGAGCCAGTGTCTTGGCCGCGCTCAATACCGGAT 365
| | | | |
QY 3299 AATACCGGCGCATAGCAGAACTTTAAAGTGTCTCATCTGAAACGTTCTTGGGG 3358
| | | | |
Db 364 AATACCGGCGCATAGCAGAACTTTAAAGTGTCTCATCTGAAACGTTCTTGGGG 305
| | | | |
QY 3359 CGAAACTCTCAAGATTTACCGCTGTGAGATCCAGTCCAGTGTAAACCACTCGTGCA 3418
| | | | |
Db 304 CGAAACTCTCAAGATTTACCGCTGTGAGATCCAGTCCAGTGTAAACCACTCGTGCA 245
| | | | |
QY 3419 CCNAAGTATCTTCAAGATCTTTTACTTTCACAGGCTTTCTGGTACAAAAACAGA 3478
| | | | |
Db 244 CCNAAGTATCTTCAAGATCTTTTACTTTCACAGGCTTTCTGGTACAAAAACAGA 185
| | | | |
QY 3479 AGGCAAAATGCCGCAAAAAAGGAATAGGCGCAGACGGAATGTGAATCTCATCTC 3538
| | | | |
Db 184 AGGCAAAATGCCGCAAAAAAGGAATAGGCGCAGACGGAATGTGAATCTCATCTC 125
| | | | |
QY 3539 TTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATTTGTCTATGAGCGGATACATA 3598
| | | | |
Db 124 TTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATTTGTCTATGAGCGGATACATA 65
| | | | |
QY 3599 TTTGAATGTATTTTGAAAAAATAACAATAGGGGTTCCGCGCATTTCCCGAAAAAGTG 3658
| | | | |
Db 64 TTTGAATGTATTTTGAAAAAATAACAATAGGGGTTCCGCGCATTTCCCGAAAAAGTG 5
| | | | |
QY 3659 CCAC 3662
| | | | |
Db 4 CCAC 1

Search completed: January 26, 2005, 17:51:53
Job time : 1991.66 sec8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:19:36 ; Search time 304.259 Seconds
(without alignments)
8554.922 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662
Sequence: 1 ctaactgttaagcgtaata.....attccccgaagtgccac 3662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2639	72.1	3681	1 US-07-924-028A-6	Sequence 6, Appli
2	2639	71.6	5314	1 US-07-924-028A-4	Sequence 4, Appli
3	2630	71.6	3699	3 US-08-646-538-6	Sequence 6, Appli
4	2620	71.6	3699	3 US-09-503-222-6	Sequence 6, Appli
5	2469	67.4	5277	4 US-09-890-848-16	Sequence 16, Appli
6	2448	66.9	3956	4 US-09-402-266B-21	Sequence 21, Appli
7	2305	62.9	4102	4 US-09-402-266B-20	Sequence 20, Appli
8	2302	62.9	4088	4 US-09-402-266B-18	Sequence 18, Appli
9	2279	62.2	4119	4 US-09-993-170-1	Sequence 1, Appli
10	2172	59.3	7102	3 US-09-138-024-20	Sequence 20, Appli
11	2172	59.3	7102	3 US-09-404-066-20	Sequence 20, Appli
12	2172	59.3	7102	4 US-09-573-322-20	Sequence 20, Appli
13	2172	59.3	7333	4 US-09-138-024-21	Sequence 21, Appli
14	2172	59.3	7333	4 US-09-404-066-21	Sequence 21, Appli
15	2172	59.3	7333	4 US-09-573-322-21	Sequence 21, Appli
16	2170	59.3	5178	2 US-08-474-169-2	Sequence 2, Appli
17	2170	59.3	6206	2 US-08-474-169-3	Sequence 3, Appli
18	2170	59.3	2973	4 US-09-402-266B-17	Sequence 17, Appli
19	2170	59.3	3789	4 US-09-075-019-8	Sequence 8, Appli
20	2170	59.3	4583	4 US-09-402-266B-19	Sequence 19, Appli
21	2170	59.3	2961	3 US-08-446-935-6	Sequence 6, Appli
22	2170	59.3	4145	3 US-08-651-472-62	Sequence 62, Appli
23	2170	59.3	4145	3 US-08-358-928-62	Sequence 62, Appli
24	2170	59.3	4277	3 US-08-651-472-63	Sequence 63, Appli
25	2170	59.3	4277	3 US-08-358-928-63	Sequence 63, Appli
26	2170	59.3	5356	3 US-08-446-935-1	Sequence 1, Appli
27	2170	59.3	5532	3 US-08-651-472-72	Sequence 72, Appli

C 28	2170	59.3	5532	3 US-08-358-928-72	Sequence 72, Appli
C 29	2170	59.3	5534	1 US-08-452-267-3	Sequence 3, Appli
C 30	2170	59.3	5534	3 US-09-123-644-3	Sequence 3, Appli
C 31	2170	59.3	6811	3 US-08-651-472-67	Sequence 67, Appli
C 32	2170	59.3	6811	3 US-08-358-928-67	Sequence 67, Appli
C 33	2170	59.3	6926	3 US-08-651-472-69	Sequence 69, Appli
C 34	2170	59.3	6926	3 US-08-358-928-69	Sequence 69, Appli
C 35	2170	59.3	7090	4 US-09-714-520-18	Sequence 18, Appli
C 36	2170	59.3	7214	4 US-09-438-268-1	Sequence 1, Appli
C 37	2170	59.3	7432	4 US-09-367-891A-6	Sequence 6, Appli
C 38	2170	59.3	7474	4 US-09-845-917A-26	Sequence 26, Appli
C 39	2170	59.3	8560	4 US-09-936-572-11	Sequence 11, Appli
C 40	2170	59.3	11233	3 US-08-980-832-27	Sequence 27, Appli
C 41	2170	59.3	11233	4 US-09-920-923B-27	Sequence 27, Appli
C 42	2169	59.2	3737	2 US-08-784-208-1	Sequence 1, Appli
C 43	2169	59.2	3796	4 US-09-470-661A-32	Sequence 32, Appli
C 44	2169	59.2	4399	4 US-09-911-781-27	Sequence 27, Appli
C 45	2169	59.2	4399	4 US-10-400-902-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1

US-07-924-028A-6
Sequence 6, Application US/07924028A

Patent No. 5470573

GENERAL INFORMATION:

APPLICANT: Lubitz Werner, Szostak, Michael P.

TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS

TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/924,028A

FILING DATE: 30-SEP-1992

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP91/00308

FILING DATE: 02-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 40 05 874

FILING DATE: 24-FEB-1990

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5470573man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: HUBR 1027

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 3681 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-07-924-028A-6

Query Match 72.1%, Score 2639, DB 1, Length 3681;
Best Local Similarity 83.5%, Pred. No. 0;
Matches 3075, Conservative 0, Mismatches 580, Indels 28, Gaps 6;

QY 3 AAATGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCACTCAT 62
Db 1 AAATTTTAAACGTATATATTTTGTAAATTCGCGTTAAATTTTGTAAATCACTCAT 60
QY 63 TTTTAAACCAATAGCGCGAAATTCGCGAAATTCCTTAAATCAAAAGATAGACCGA 122
Db 61 TTTTAAACCAATAGCGCGAAATTCGCGAAATTCCTTAAATCAAAAGATAGACCGA 120
QY 123 TAGGGTTAGTGTGTTCCAGTTTGGAAACAAGATCACTAATTAAGAAAGTGACTCA 182
Db 121 TAGGGTTAGTGTGTTCCAGTTTGGAAACAAGATCACTAATTAAGAAAGTGACTCA 180
QY 183 ACGTCAAAAGGCGGAAAAACCGTCTATACAGGAGATGGCCCACTAGGTGAACCATCA 242
Db 181 ACGTCAAAAGGCGGAAAAACCGTCTATACAGGAGATGGCCCACTAGGTGAACCATCA 240
QY 243 AATCAAGTTTGTGGGGTGAAGGTGCGCTTAAGCACTAATTCGAAACCTTAAGAGGCC 302
Db 241 AATCAAGTTTGTGGGGTGAAGGTGCGCTTAAGCACTAATTCGAAACCTTAAGAGGCC 300
QY 303 CCGAATTAGACTTGAACGAGGAAAGCGCGCAAGCTGGCGAAGAAAGGAAAGAAAG 362
Db 301 CCGAATTAGACTTGAACGAGGAAAGCGCGCAAGCTGGCGAAGAAAGGAAAGAAAG 360
QY 363 CGAAAGAGCGGCGCGCTAGAGGCGCTGAGCAAGTGAAGCGTCAAGCTGGCGAAGCA 422
Db 361 CGAAAGAGCGGCGCGCTAGAGGCGCTGAGCAAGTGAAGCGTCAAGCTGGCGAAGCA 420
QY 423 CACCGCGCGCGTTAATGCGCGCTTACAGGCGCGCTCCATTTGCGCAATTCAGGTCGCA 482
Db 421 CACCGCGCGCGTTAATGCGCGCTTACAGGCGCGCTCCATTTGCGCAATTCAGGTCGCA 480
QY 483 ACTGTGGGAAAGGCGCATCGGTGCGGCGCTTCCGTTAATTCGCAAGCTGGCGAAGG 542
Db 481 ACTGTGGGAAAGGCGCATCGGTGCGGCGCTTCCGTTAATTCGCAAGCTGGCGAAGG 540
QY 543 GATGTGCTGCAAGCGCATTAATTTGGGTAACGCAAGGTTTCCAGTCAAGAGCTGTA 602
Db 541 GATGTGCTGCAAGCGCATTAATTTGGGTAACGCAAGGTTTCCAGTCAAGAGCTGTA 600
QY 603 AAAGGAGCGCGCAATGAGCGCGCTCGTTCAATTCAGCTTTGAAACCGCTGAGGACGG 662
Db 601 AAAGGAGCGCGCAATGAGCTGTAATGAGCTCACTAAGGCGAAATGAGCTCCACCGG 660
QY 663 AGACTGCGCGTGAATGCTTTACAGGCTGATGAGCAAGATGCTGAGACCG 722
Db 661 GTGGCGCGCTTGAATGCTGACTCTCAATCTGCTGATGCTGATGCTGATGCTG 720
QY 723 AGCGAAGCAAGCACTAATTAACCTTGAAGAAATATCAATTTGATGAGCTGATGAG 782
Db 721 AGCGAAGCAAGCACTAATTAACCTTGAAGAAATATCAATTTGATGAGCTGATGAG 780
QY 783 AATATCATGCGTAAATTAAGCAATGAGATCTGTAATACGACTCACTAAGGCGAAATG 842
Db 781 GCGAACAACCGCTACCGGCTCTGACGCGCTGTCTGCTCCGCACTCCGCTTAAGCA 840
QY 843 GGTACCGGCGCGCGCTCTGAGGTGACGATGCA-----TAACTGATATGCAATTC 896
Db 841 AGCTGACCGCTCTCCGAGCTGACATGCTCAGAGATTTTCAACGTATCAACGAAACG 900
QY 897 CTGACCGCGGAGATCACTAGTCTAAGGCGCGCAACGCGGCTGAGACTCCAGCTT 956
Db 901 CGGAGGCAATGAGTGTGATGCTTTGTGCAATTCGTCTTAAGTAAGCAATTCG 960
QY 957 TTGTTCCTTTAGTGAAGGTTAATTAAGTCCCAT-----GCGTCAATTTTACGACAC 1009
Db 961 TAAAGTGTGACCTGCGGATCAACGCTTCCAGTACGACAGAAAGCAATGATGTTAA 1020
QY 1010 TATCTTTTAGGGTTAATCTAGCTGATCAGAGATCAATGCTGGGCTTTTTCGGCT 1069
Db 1021 TTTTCGAGAAAGATTCGAGAAAGATCAATTAAGATTTGAATCTTTGTGTCT 1080
QY 1070 CAGTATGCGCCAACTGCGCTATCTGGGCACTGGGAGGAAAGAACCCGTCCTTTTC 1129

Db 1081 TCACATGCGTAAATCTCATGTTTGAATGCGCTTGAAGATCCGCGCAACCTTGCAATG 1140
QY 1130 CCGCGAGGTTGAAGGCGCATGAAAGATTTGCGAGATGATGCTGCTGCAATTAAGT 1189
Db 1141 TGAAGTGAATCTTGAAGATCCCGACGCTCGACGCAATTAATTAATTTCCGTAAT 1200
QY 1190 TGAAGGAAAGCGACGTTTACATGATGATTCGAGAAAGGTGTCATGACGCTTTAA 1249
Db 1201 TGAAGGCTTCATGATGAGACAGGCGCTT--GAATGTGACGAGATGACATTAATAG 1258
QY 1250 CCGTGAATGTTGCTTGAAGCGACCTGAGATACCAATTCGCGGCTTTTCGAGACA 1309
Db 1259 CAATGAGGAGCAATTAATCAACAG--AGCAGAAAGGAGGATATCCCAAGT 1315
QY 1310 GTTCGAGATGTCAGCCCGAAGCGCATCAAGCAACCGAAACAATACCGCGACAG----- 1363
Db 1316 CCAAGTAAACATTAACGCAAGCTTCAAGCAGACGACAGAGAGCGTCACTAGCA 1375
QY 1364 CCGAATGCTGCGCGGATGACATTAATGACAGCGGTGCGCGCTGAGATTAATGCT 1423
Db 1376 TCCAACTTTGTTACTCTGACAGAAATGAATCATCTTCGTTAAATTCAAACGCGAG 1435
QY 1424 CAGGAGAGCGGCTAT-----CTGCTGATGACCGCAAAATGAGACATGATACCCG 1479
Db 1436 AAGCTGAATGAAATTCGACCTGAGGCGGCGCGCTTCCAGCTTTTGTCTTTA 1495
QY 1480 AGTTACCGCGCGCGCGCTTGGGTTAATGATGATGATGATGATGATGATGATGATG 1539
Db 1496 GTGAGGTTTAAATTCGAGCTTGGGTTAATGATGATGATGATGATGATGATGATG 1555
QY 1540 TTATCGCTCAATTTTCAACAACAATACAGCGGAGACATAAGTGAAGCTTGCGG 1599
Db 1556 TTATCGCTCAATTTTCAACAACAATACAGCGGAGACATAAGTGAAGCTTGCGG 1615
QY 1600 TGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1675
Db 1616 TGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1719
QY 1660 GGGAAACCTGCTGACGCTGATTAATGAATGCGGCCAAGCGCGGAGAGCGGCTT 1735
Db 1676 GGGAAACCTGCTGACGCTGATTAATGAATGCGGCCAAGCGCGGAGAGCGGCTT 1779
QY 1720 GCGTATTTGGCGCTTTCGCTTCTCGCTCACTGACCTGCTGCTGCTGCTGCTG 1795
Db 1736 GCGTATTTGGCGCTTTCGCTTCTCGCTCACTGACCTGCTGCTGCTGCTGCTG 1839
QY 1780 GCGGCGAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1855
Db 1796 GCGGCGAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1899
QY 1840 TAAAGCAAGAAAGAAATGATGAGCAAAAGCGCAAAAGGCGCAAGAACCTGTAAGAG 1959
Db 1856 TAAAGCAAGAAAGAAATGATGAGCAAAAGCGCAAAAGGCGCAAGAACCTGTAAGAG 1975
QY 1900 CCGCTTGTGCGCTTTTTCATAGGCTCGCGCCCTTGAAGCAATCAAAATTCAGAG 1995
Db 1916 CCGCTTGTGCGCTTTTTCATAGGCTCGCGCCCTTGAAGCAATCAAAATTCAGAG 2019
QY 1960 CTCAAGTCAAGATGAGGAAAGCGCAAGCAAGCAATTAAGATTAAGAGGCTTCCCGT 2035
Db 1976 CTCAAGTCAAGATGAGGAAAGCGCAAGCAAGCAATTAAGATTAAGAGGCTTCCCGT 2079
QY 2020 AAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2095
Db 2036 AAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2139
QY 2080 TCTCCCTTGGGAAAGCGGCTTCTCATAGCTCAAGCTGATGATGATGATGATGATG 2155
Db 2096 TCTCCCTTGGGAAAGCGGCTTCTCATAGCTCAAGCTGATGATGATGATGATGATG 2199
QY 2140 GTAGGCTGCTGCTCAAGCTGAGCTGATGAGCAAAACCCCGTTCAAGCCGACGCTG

Db 2156 GTAGTCGTTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTG 2215
Qy 2200 CGCTTATCCGGTAACCTATGCTTTGAGTCCAAACCCGGTAAGACGACCTTATGGCCACT 2259
Db 2216 CGCTTATCCGGTAACCTATGCTTTGAGTCCAAACCCGGTAAGACGACCTTATGGCCACT 2275
Qy 2260 GGCAGCAGCAGCTGTTAAACAGATTAGCAGACGAGATTTAGGCGGCTGTACAGACTT 2319
Db 2276 GGCAGCAGCAGCTGTTAAACAGATTAGCAGACGAGATTTAGGCGGCTGTACAGACTT 2335
Qy 2320 CTGTAAGTGTGGCTTAACCTACCGCTACCTAGAAAGACAGTATTTGGTATCTGGCTCT 2379
Db 2336 CTGTAAGTGTGGCTTAACCTACCGCTACCTAGAAAGACAGTATTTGGTATCTGGCTCT 2395
Qy 2380 GCTAAGCAGTACCTTCGGAAGAAAGTTGTAAGCTTTGATCTGGCAACCAACAC 2439
Db 2396 GCTAAGCAGTACCTTCGGAAGAAAGTTGTAAGCTTTGATCTGGCAACCAACAC 2455
Qy 2440 CGCTGTAGCGGTGGTGTGTTGTTGCAAGCAGAGATTACGGCGAGAAAGAAAGATC 2499
Db 2456 CGCTGTAGCGGTGGTGTGTTGTTGCAAGCAGAGATTACGGCGAGAAAGAAAGATC 2515
Qy 2500 TCAAGAGATCTCTTGAATCTTTCTACCGGGCTGTACGCTGACGTGAAGAAACCTCAG 2559
Db 2516 TCAAGAGATCTCTTGAATCTTTCTACCGGGCTGTACGCTGACGTGAAGAAACCTCAG 2575
Qy 2560 TTAAAGGATTTTGTCTATGAGATTATCAAAAGATCTTCACTAGATCTTTTAAATTA 2619
Db 2576 TTAAAGGATTTTGTCTATGAGATTATCAAAAGATCTTCACTAGATCTTTTAAATTA 2635
Qy 2620 AAAATGAAGTTTAAATCAATCTAAGATATATGAGATAAAGCTTGTGACATTTTCA 2679
Db 2636 AAAATGAAGTTTAAATCAATCTAAGATATATGAGATAAAGCTTGTGACATTTTCA 2695
Qy 2680 ATGCTTATCAGTAGAGGACCTATCTACGCGATCTGTCTATTTGTTTATCATAGTTGC 2739
Db 2696 ATGCTTATCAGTAGAGGACCTATCTACGCGATCTGTCTATTTGTTTATCATAGTTGC 2755
Qy 2740 CTGACTCCCGCTGCTGTATGATTAATCAGATACGGAAGGCTTACCATCTGCCCCAGTGC 2799
Db 2756 CTGACTCCCGCTGCTGTATGATTAATCAGATACGGAAGGCTTACCATCTGCCCCAGTGC 2815
Qy 2800 TGCAATATACCGGAGAGCCACGCTACCGGCTCCAGATTATCAGAAATTAACCGACC 2859
Db 2816 TGCAATATACCGGAGAGCCACGCTACCGGCTCCAGATTATCAGAAATTAACCGACC 2875
Qy 2860 AGCCGAAAGGCGCAGGCGCAGAAAGTGTCTGTCAACTTATCCGCTTCATCCAGTCTAT 2919
Db 2876 AGCCGAAAGGCGCAGGCGCAGAAAGTGTCTGTCAACTTATCCGCTTCATCCAGTCTAT 2935
Qy 2920 TAAATGTTGCGGGAAGCTAGAGTAAGTATGCGCAGTTAATAGTTTGCAGAACGTTGT 2979
Db 2936 TAAATGTTGCGGGAAGCTAGAGTAAGTATGCGCAGTTAATAGTTTGCAGAACGTTGT 2995
Qy 2980 TGCAATGCTACAGAGCAATCGTGTGTCACGCTCGTGGTGTATGCTTATCATCAGTCTC 3039
Db 2996 TGCAATGCTACAGAGCAATCGTGTGTCACGCTCGTGGTGTATGCTTATCATCAGTCTC 3055
Qy 3040 CGGTTTCCCAAGATCAAGGCGAGTATGATATCCCATGTTTGTGCAAAAGGCGTTAG 3099
Db 3056 CGGTTTCCCAAGATCAAGGCGAGTATGATATCCCATGTTTGTGCAAAAGGCGTTAG 3115
Qy 3100 CTCTTTGGTCTCTCGATCGTTGTCAAGAGTAAGTTGGCGCAGTGTATCATCATGCT 3159
Db 3116 CTCTTTGGTCTCTCGATCGTTGTCAAGAGTAAGTTGGCGCAGTGTATCATCATGCT 3175
Qy 3160 TATGGCAGCATGCTGATTAATCTCTTACGTATGCAATCCGTAAGGCTTTTCTGTGAC 3219
Db 3176 TATGGCAGCATGCTGATTAATCTCTTACGTATGCAATCCGTAAGGCTTTTCTGTGAC 3235
Qy 3220 TGGTGAATCTCAACCAAGTCAATCTGAGATAGTATGAGGCGACGAGTTGCTCTTG 3279
Db 3236 TGGTGAATCTCAACCAAGTCAATCTGAGATAGTATGAGGCGACGAGTTGCTCTTG 3295

Qy 3280 CCGCGCTCAATACGGGATTAATACCGGCGCCATAGCAGAACTTTAAAGTGTCTATCAT 3339
Db 3396 CCGCGCTCAATACGGGATTAATACCGGCGCCATAGCAGAACTTTAAAGTGTCTATCAT 3355
Qy 3340 TGAAGAAAGTCTCTCGGGGCGAAAGCTCTCAAGGATCTTACCGCTGTGAGTCAAGTTC 3399
Db 3356 TGAAGAAAGTCTCTCGGGGCGAAAGCTCTCAAGGATCTTACCGCTGTGAGTCAAGTTC 3415
Qy 3400 GATGTAACCACTGTCACCCCACTGATCTTTCAGATCTTTTACCTTTCACAGGTTTC 3459
Db 3416 GATGTAACCACTGTCACCCCACTGATCTTTCAGATCTTTTACCTTTCACAGGTTTC 3475
Qy 3460 TGGGTGAGCAAAAACAGAGGCAAAATGCCGCAAAAAGGAAATTAAGGCGACAGGAA 3519
Db 3476 TGGGTGAGCAAAAACAGAGGCAAAATGCCGCAAAAAGGAAATTAAGGCGACAGGAA 3535
Qy 3520 ATGTTGAATCTCACTACTCTCTCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTG 3579
Db 3536 ATGTTGAATCTCACTACTCTCTCTTTTCAATATTAATGAAGCATTTATCAGGGTTATTG 3595
Qy 3580 TCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGCG 3639
Db 3596 TCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGCG 3655
Qy 3640 CACATTTCCCGGAAAGTCCAC 3662
Db 3656 CACATTTCCCGGAAAGTCCAC 3678

RESULT 2
US-07-924-028A-4
; Sequence 4, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pelfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-9884
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
US-07-924-028A-4Query Match 71.6%; Score 2623.6; DB 1; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3053; Conservative 0; Mismatches 579; Indels 28; Gaps 6;

QY 20 ATTTGTTAAATTCGCGCTTAAATTTTGTATTAATCAGTCATTTTAAACCAATAGGCC 79

Db 1651 ATTTGTTAAATTCGCGCTTAAATTTTGTATTAATCAGTCATTTTAAACCAATAGGCC 79

QY 80 GAAATCGGCAAAATTCCTTATTAATCAAAAGAAATAGACGAGATAGGGTTAGTGTGT 1710

Db 1711 GAAATCGGCAAAATTCCTTATTAATCAAAAGAAATAGACGAGATAGGGTTAGTGTGT 139

QY 140 CCAAGTTTGAACAAGATCCACTATTAAAGACGTGACTCCAGCTCAAGGAGGCGAATA 1770

Db 1771 CCAAGTTTGAACAAGATCCACTATTAAAGACGTGACTCCAGCTCAAGGAGGCGAATA 199

QY 200 ACCGTCATCAGGCGCATGCGCCCACTACGTGAACCATCACCTTATCAAGTTTGTGGG 1830

Db 1831 ACCGTCATCAGGCGCATGCGCCCACTACGTGAACCATCACCTTATCAAGTTTGTGGG 259

QY 260 TCGAGGTGCGGTAAAGCACTAAATCGAAACCTTAAAGGAGCCCGCATTTAGAGCTTGA 1890

Db 1891 TCGAGGTGCGGTAAAGCACTAAATCGAAACCTTAAAGGAGCCCGCATTTAGAGCTTGA 319

QY 320 CCGGGAAGCCCGGCAACGTGCGCAAGAAAGGAAAGGAAAGGAAAGGAGGCGCGCT 1950

Db 1951 CCGGGAAGCCCGGCAACGTGCGCAAGAAAGGAAAGGAAAGGAAAGGAGGCGCGCT 379

QY 380 AGGGGCGTTCGCAAGTGAAGCGGTCAAGCTGCGGTAAACCAACACCGCCCGCGCTTAT 2010

Db 2011 AGGGGCGTTCGCAAGTGAAGCGGTCAAGCTGCGGTAAACCAACACCGCCCGCGCTTAT 439

QY 440 GCGCGCTACAGGCGCGGTCCCAATGCGCATTCAGGCTGCGCAATGTTGGGAAGGCGA 2070

Db 2071 GCGCGCTACAGGCGCGGTCCCAATGCGCATTCAGGCTGCGCAATGTTGGGAAGGCGA 499

QY 500 TCGGTGCGGCGCTTTCGCTATTAACGACGTCGCGGCAAGGAGGATGTGCTGCAAGCGA 2130

Db 2131 TCGGTGCGGCGCTTTCGCTATTAACGACGTCGCGGCAAGGAGGATGTGCTGCAAGCGA 559

QY 560 TTAAAGTTGAGTAAACGCAAGGTTTCCAGTCAAGACGTTGTAAACGACGCGCAGTGA 2190

Db 2191 TTAAAGTTGAGTAAACGCAAGGTTTCCAGTCAAGACGTTGTAAACGACGCGCAGTGA 619

QY 620 CCGCGCTCGTTCATTCAGGTTTGAACCGGTGAGGACGCGGAGACGCTGCGGTCAAT 2250

Db 2251 TTGTATACGACTCATATAGGCGCAATTGGAAGCTCCACCGCGGTGCGGCTCTTGT 679

QY 680 GTGTTTACAGGTGATGAGACAGATGAAAGTGTGACACGCTGCAACACCGCTTGT 2310

Db 2311 ATGTGCACTCTCAATCAATCTGCTCTGATGCGCAATGTTAAACGAGTATATACACT 739

QY 740 GATTAAACCTAGAAAGATTAATCAATTTGTGACGTAAAGGATTAATCATGCGTAAAT 2370

Db 2371 CGCTATGCTACGAGTATGAGTATGACATCACTATAGGCGCAATTTGGGTACCGGCGT 799

QY 800 TGACGATGAGATTTGTATATACATCACTATAGGCGCAATTTGGGTACCGGCGT 2430

Db 2431 CGCCTACGAGGCTTGTCTGCTCCGCGCATCCGCTTACAGACAAAGCTGTACCGTCCG 859

QY 860 CGAGGTGACGATATGCA-----TAACTGTATATGCAATTTCCGTGACGCGGAGATC 2490

Db 2491 GAGAGTCAATGTGTCAAGGTTTTCACCGTCAATCAACGAAACGCGAGGCAATAGGTC 913

QY 914 CACTAGTTTCAAGGCGCGCAACGCGGTGAGCTCAGCTTTGTGTCTTTAGTGAAG 2550

Db 2551 GAGTGTGTTGTGAGCAATTCCTTAAAGTGAAGCAATTCCTTAAAGTGTGACTGTGC 973

QY 974 GGTATTAATGATCCAT-----GCGTCAATTTTACGAGACTATCTTTTAAAGGTTAA 1026

Db 1026 GGTATTAATGATCCAT-----GCGTCAATTTTACGAGACTATCTTTTAAAGGTTAA 1026

Db

QY 2611 GATCAACCGCTTCAGTAGACCAAGAACAAATGATTGTAAATTTGAGAGAAATGCG 2670

Db 1027 TCTAGCTGATAGAGATCATATTCGTGCGGTCTTTTTCGCGCTCATCTATGCCAAGCT 1086

QY 2671 CAGAGAAAGTCAATCATTAAGAGTTGAACCTCTTTGTTGTTGTCGACATGGTATCCT 2730

Db 1087 GCGCTATCTGCGCATCGGGAGGAAAGAACCCGCTTTCGCCGAGTTGAAGCGG 1146

QY 2731 CATGTTTAAATGCGCTTAGAGATCCGCAAGCTTGACATGCTCAGGTGACCTTGA 2790

Db 1147 CATGAAAGATTTGCCGAGATGACTGCTGCTGCAATTAAGCTTGAAGGAAAGGACGCT 1206

QY 2791 GATGCCGAGAGCTTCAGCGCATTAATATGTTTCCGTAATTCAGGCGCTTCATGAT 2850

Db 1207 TTACCATGATGATTTCGGAAGGTGTGCGCATGACAGCTTAAAGGTGAATGTTCTTC 1266

QY 2851 GAGACAGGCGGTTT--GAATGTTGACGGATGACATATATAGCAATGACGCGCAATA 2908

Db 1267 AGGCACTGCGGATACCAATTTGCTGCGGCTTTTCGAGACAGATTCGAGATGTTAGCG 1326

QY 2909 AACTCAACAG--AGCAGAAAGGAGGATATCCCAAAAGTCAAGCTTACATTAACG 2965

Db 1327 CGAAGGCGATACAGCAACCGCAACATACCGGCAAG-----CGGAAGTCCGTCGCG 1380

QY 2966 CAAGCTTCAAGCGACGACGAGCAGAGAGGCGTCACTAGACCAATCCAACTTTGTACTC 3025

Db 1381 GTGTCAATTAATGACAGCGGTGCGGCGCTGGATTTATTCGTACAGCGAGACGCGGTAT 1439

QY 3026 GTCAAAATATGAATCATCTTGGTTAAATCCAAACCGCAGAAAGCTGAAATGAAAT 3085

Db 1440 ---CTGTGCTGAGTCCGCGCAAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1496

QY 3086 GACCTGAGAGGCGGCGGCTGACCGAGCTTTGTCTTGTAGAGGTTAATTCGA 3145

Db 1497 GCTTGGCTATCATGATCATATGCTGTTTCCGTGTTAAATTTGTATTCGTCACAAATTC 1556

QY 3146 GCTTGGCTATCATGATCATATGCTGTTTCCGTGTTAAATTTGTATTCGTCACAAATTC 3205

Db 1557 CACACAAATACAGCGCGGCAACATTAAGTTAAAGCTGCGGCTTAATGATGATGAT 1616

QY 3206 CACACAAATACAGCGCGGCAACATTAAGTTAAAGCTGCGGCTTAATGATGATGAT 3265

Db 1617 AACTCATTTAATTCGTTGCTGCTCACTGCGGCTTTCAGTGGGAAACCTGTGTCG 1676

QY 3266 AACTCATTTAATTCGTTGCTGCTCACTGCGGCTTTCAGTGGGAAACCTGTGTCG 3325

Db 1677 AACTCATTTAATTCGTTGCTGCTCACTGCGGCTTTCAGTGGGAAACCTGTGTCG 3385

QY 3386 AACTCATTTAATTCGTTGCTGCTCACTGCGGCTTTCAGTGGGAAACCTGTGTCG 3445

Db 1737 CCGCTTCCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1796

QY 3445 CCGCTTCCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445

Db 1797 CCGCTTCCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445

QY 3446 CCGCTTCCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3505

Db 1857 TGTAGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGAGGATGAGCAAAAGCA 3505

QY 3506 TGTAGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGAGGATGAGCAAAAGCA 3565

Db 1917 TGTAGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGAGGATGAGCAAAAGCA 3565

QY 1917 TGTAGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGAGGATGAGCAAAAGCA 3565

Db 3566 TGTAGCAAAAGGCGAGCAAAAGGCGAGAAACGTTAAAGGCGAGGATGAGCAAAAGCA 3625

QY 1977 GAAACCGCAGAGCAATTAAGATTCAGAGGCTTTTCCCTGGAAGCTCCCTGCGGCT 2036

Db 3626 GAAACCGCAGAGCAATTAAGATTCAGAGGCTTTTCCCTGGAAGCTCCCTGCGGCT 3685

QY 2037 GTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGCG 2096

Db 3686 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTCCGGAAGCG 3745

2097 TGGCGCTTTCATAGCTCAGCGCTAGGTATCTCAGTTCCGTTAGTGGTTCCTCCTCA 2156
 3146 TGGCGCTTTCATAGCTCAGCGCTAGGTATCTCAGTTCCGTTAGTGGTTCCTCCTCA 3805
 2157 AGCTGGGCTGTGTGACGACACCCCGTTACGCCGCGCGGCGCTTATCCGGTAATCT 2216
 3806 AGCTGGGCTGTGTGACGACACCCCGTTACGCCGCGCGGCGCTTATCCGGTAATCT 3865
 2217 ATCGCTTGAAGTCCACCCCGGTAGACAGACTTATCCGCACTGGCAGAGCACTGGTA 2276
 3866 ATCGCTTGAAGTCCACCCCGGTAGACAGACTTATCCGCACTGGCAGAGCACTGGTA 3925
 2277 ACAGGATTAGCAGACGAGGTATGTAGGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTA 2336
 3926 ACAGGATTAGCAGACGAGGTATGTAGGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTA 3985
 2337 ACTACGGCTACCTAGAGGACAGATTTGGTATCTGGCGCTGTGCTGAAGCCAGTTACT 2396
 3986 ACTACGGCTACCTAGAGGACAGATTTGGTATCTGGCGCTGTGCTGAAGCCAGTTACT 4045
 2397 TCGGAAAAAGATTGGTAGCTTTGATCCGCGCAACAAACACCGCTGGTAGCGGTGGT 2456
 4046 TCGGAAAAAGATTGGTAGCTTTGATCCGCGCAACAAACACCGCTGGTAGCGGTGGT 4105
 2457 TTTTGTGTTCAGACGACGATTAACGCGCAAAAAAGATCTCAAGAGATCTTTGA 2516
 4106 TTTTGTGTTCAGACGACGATTAACGCGCAAAAAAGATCTCAAGAGATCTTTGA 4165
 2517 TCTTTTACGAGGCTGACGCTCAGTGAACGAAATCTCAGTTAAGGATTTGGTCA 2576
 4166 TCTTTTACGAGGCTGACGCTCAGTGAACGAAATCTCAGTTAAGGATTTGGTCA 4225
 2577 TGAATTAATCAAAAAGATCTTACCTAGATCTTTTAAATTAATTAATTAATTAAT 2636
 4226 TGAATTAATCAAAAAGATCTTACCTAGATCTTTTAAATTAATTAATTAATTAAT 4285
 2637 CAATCTAAGTATATATAGATTAATTTGGTCTGACAGATTAACCAATGCTTAATCAATGAGG 2696
 4286 CAATCTAAGTATATATAGATTAATTTGGTCTGACAGATTAACCAATGCTTAATCAATGAGG 4345
 2697 CACCTATCTACGCGATCTGTCTATTTGCTTCAATCCATGTTGCTCCCGCGTGGT 2756
 4346 CACCTATCTACGCGATCTGTCTATTTGCTTCAATCCATGTTGCTCCCGCGTGGT 4405
 2757 AGATACTACGATACGAGGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGCGAG 2816
 4406 AGATACTACGATACGAGGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGCGAG 4465
 2817 ACCCAGCTCACCAGCTCCAGATTTATCAGCAATTAACAGCCAGCCGGAAGGCGGAGC 2876
 4466 ACCCAGCTCACCAGCTCCAGATTTATCAGCAATTAACAGCCAGCCGGAAGGCGGAGC 4525
 2877 GCAGAGGTGCTCTGCACTTATCCGCTCATATCAGTCTTAATTTGTTGCGGAGAG 2936
 4526 GCAGAGGTGCTCTGCACTTATCCGCTCATATCAGTCTTAATTTGTTGCGGAGAG 4585
 2937 CTAGAGTAAGTATGCTCCAGTATTAATTTGCGCAAGTGTGCTGCAATGCTTACAGGCA 2996
 4586 CTAGAGTAAGTATGCTCCAGTATTAATTTGCGCAAGTGTGCTGCAATGCTTACAGGCA 4645
 2997 TCGTGGTGTCAAGCTCGTGTGTTGGTATGCTTCACTAGCTCCGGTTCGCAAGATCAA 3056
 4646 TCGTGGTGTCAAGCTCGTGTGTTGGTATGCTTCACTAGCTCCGGTTCGCAAGATCAA 4705
 3057 GCGGAGTTACATGATCCCGCATGTTGTCAGAAAAAGGAGTTAGCTTCCGCTCCGGA 3116
 4706 GCGGAGTTACATGATCCCGCATGTTGTCAGAAAAAGGAGTTAGCTTCCGCTCCGGA 4765
 3117 TCGTTGTCAAGTAAGTTGGCCGAGTGTATCACTAGTTATGCGAGCATGTCATA 3176
 4766 TCGTTGTCAAGTAAGTTGGCCGAGTGTATCACTAGTTATGCGAGCATGTCATA 4825

3177 ATTCCTTACGTGATGCGCATCCGTAAGATGCTTTCTGTGACGTGGTAGTACCAACA 3236
 4826 ATTCCTTACGTGATGCGCATCCGTAAGATGCTTTCTGTGACGTGGTAGTACCAACA 4885
 3237 AGTCAATCTGAGAAATAGTATAGCGGACGAGTGTGCTTTGCGCGCTCAATACGAG 3296
 4886 AGTCAATCTGAGAAATAGTATAGCGGACGAGTGTGCTTTGCGCGCTCAATACGAG 4945
 3297 ATTAATACCGCGCATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTGG 3356
 4946 ATTAATACCGCGCATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTGG 5005
 3357 GCGGAAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTCCATTTGCCACTTGG 3416
 5006 GCGGAAAACTCTCAGAGATCTTACCGCTGTGAGATCCAGTTCCATTTGCCACTTGG 5065
 3417 CACCAACTGATCTTACGATCTTTTACCTTACCAAGCGTTTCTGGTAGAGAAAAAG 3476
 5066 CACCAACTGATCTTACGATCTTTTACCTTACCAAGCGTTTCTGGTAGAGAAAAAG 5125
 3477 GAAGCAAAATGCGCAAAAAAGGAAATAGCGCGACACGAAATGTTGAATCTCATAC 3536
 5126 GAAGCAAAATGCGCAAAAAAGGAAATAGCGCGACACGAAATGTTGAATCTCATAC 5185
 3537 TCTTCCTTTTCAATATATTAAGCATTTATCAGGTTATTTGCTCATAGCGGATACA 3596
 5186 TCTTCCTTTTCAATATATTAAGCATTTATCAGGTTATTTGCTCATAGCGGATACA 5245
 3597 TATTGAATGATTTAGAAAAATAAACAATAGGAGTTCCGCGCACATTTCCCGGAAAG 3656
 5246 TATTGAATGATTTAGAAAAATAAACAATAGGAGTTCCGCGCACATTTCCCGGAAAG 5305
 3657 TGCCAC 3662.
 5306 TGCCAC 5311

RESULT 3
 US-08-646-538-6
 / Sequence 6, Application US/08646538
 / Patent No. 6027881
 / GENERAL INFORMATION:
 / APPLICANT: Pavlakis, George N.
 / APPLICANT: Gaetanaris, George A.
 / APPLICANT: Stauber, Roland H.
 / APPLICANT: Vournakis, John N.
 / TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
 / TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
 / NUMBER OF SEQUENCES: 37
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Townsend and Townsend and Crew LLP
 / STREET: Two Embarcadero Center, 8th Floor
 / CITY: San Francisco
 / STATE: California
 / COUNTRY: USA
 / ZIP: 94111-3834
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: PatentIn Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/646,538
 / FILING DATE: No. 6027881 yet assigned
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Weber, Kenneth A.
 / REGISTRATION NUMBER: 31,677
 / REFERENCE/DOCKET NUMBER: 015280-249000
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (415) 576-0200
 / TELEFAX: (415) 576-0300
 / INFORMATION FOR SEQ ID NO: 6:

Query Match	71.6%;	Score 2620.6;	DB 3;	Length 3699;
Best Local Similarity	83.3%;	Pred. No. 0;		
Matches 3078;	Conservative			

QY	3	AAATTGTAAGCGTTAATAATTTTCTTTT	Indels	40;	Gaps	7;

Db	3	AAATGTGAAACGTTAATAATTGTTTAAATTCGGGTTAAATTTTGTGTTAAATCACTCAT	62
Qy	63	TTTTTAACCAATAGGGCGGAATTCGGCAAAATCCCTTAATTAATCAAAAAGAAATGACCGAGA	122
Db	63	TTTTTAAACAATAGCCCGAAATCGGCAAAATCCCTTAATTAATCAAAAAGAAATGACCGAGA	122
Qy	123	TAGGGTTAGTGTGTTCCAGTTTGGAAACAAGAGTCCATATTAAGAAGCGTGAATCCA	182
Db	123	TAGGGTTAGTGTGTTCCAGTTTGGAAACAAGAGTCCATATTAAGAAGCGTGAATCCA	182
Qy	183	ACGTCAAAAGGCGAAAAACCGTATATAGGGCGATGGCCCATAGTAAACATCAACCTT	242
Db	183	ACGTCAAAAGGCGAAAAACCGTATATAGGGCGATGGCCCATAGTAAACATCAACCTT	242
Qy	243	AATCAAGTTTTTGGGGTTCGAGGTGCCTGAAGCACTTAATTCGAAACCTTAAAGGAGCC	302
Db	243	AATCAAGTTTTTGGGGTTCGAGGTGCCTGAAGCACTTAATTCGAAACCTTAAAGGAGCC	302
Qy	303	CCGATTTAGACTTGAACGGGGAAAGCCGGGCAACGTGGCGAGAAAGGAAAGGAAAG	362
Db	303	CCGATTTAGACTTGAACGGGGAAAGCCGGGCAACGTGGCGAGAAAGGAAAGGAAAG	362
Qy	363	CGAAAGGAGGGGCGCTAGAGGCGCTGAGAGGCTGACGCTGCCTGAACACCA	422
Db	363	CGAAAGGAGGGGCGCTAGAGGCGCTGAGAGGCTGACGCTGCCTGAACACCA	422
Qy	423	CACCGCGCGCTAATGCGCCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGGTCG	482
Db	423	CACCGCGCGCTAATGCGCCGCTACAGGGCGCGTC---CCATTGCGCATTCAGGGTCG	482
Qy	483	GCAACTGTGGGGAAGGCGATCGGTGGGCGCTCTTGCCTATTAAGCCAGCTGGCGAAAG	539
Db	483	GCAACTGTGGGGAAGGCGATCGGTGGGCGCTCTTGCCTATTAAGCCAGCTGGCGAAAG	539
Qy	540	GGGAGTGTGTCGAAGGCGATTAAAGTGGGTAAAGCCAGGTTTTCCAGTCAAGACGTT	599
Db	540	GGGAGTGTGTCGAAGGCGATTAAAGTGGGTAAAGCCAGGTTTTCCAGTCAAGACGTT	599
Qy	600	GTAAGGATGTCGCAAGGCGATTAAAGTGGGTAAAGCCAGGTTTTCCAGTCAAGACGTT	659
Db	600	GTAAGGATGTCGCAAGGCGATTAAAGTGGGTAAAGCCAGGTTTTCCAGTCAAGACGTT	659
Qy	660	GGCAGACTCGCGGTGCAAAATGTTTTACACCGTATGAGGAGCATGAGATGCTCGCA	719
Db	660	GGCAGACTCGCGGTGCAAAATGTTTTACACCGTATGAGGAGCATGAGATGCTCGCA	719
Qy	720	CGCTGCA---GAACACGACGCTAAGTTAACCTTAGAAAGTAATGATTTGTGACGTAC	780
Db	720	CGCTGCA---GAACACGACGCTAAGTTAACCTTAGAAAGTAATGATTTGTGACGTAC	780
Qy	777	TTAAAGATATATCGTCGTAAATTTGACGATGAGATCTGTAAACGATCACTAATAGGG	836
Db	781	TTTCGTTGGGATCTTTCCAAAGGGGAGATTTGTGTGACAGGTAA-ATGGTGTCTGTGAAA	896
Qy	837	GAATTGGGTACCGGGCCCCCTTCGAGGTGCAACGGTATCGATTAGCTTGATTCGAAATTC	896

[illegible]

Db 1920 AACGTAAGAGCGCGGTGTGGCTTTTTCATAGAGTCCGCCCTCGAGAGCAT 1979
 QY 1946 CACAAAATGACGCTCAAGTCAGAGGTGGGAAAACCCGACAGACTTAAAGATACAG 2005
 Db 1980 CACAAAATCAGCGTCAAGTCAGAGGTGGGAAAACCCGACAGACTTAAAGATACAG 2039
 QY 2006 GCGTTTCCCGTGGAAAGCTCCCTGTGGGCTCTCCGTGTCGAGACCCTGCGTTACCGGA 2065
 Db 2040 GCGTTTCCCGTGGAAAGCTCCCTGTGGGCTCTCCGTGTCGAGACCCTGCGTTACCGGA 2099
 QY 2066 TACCTGTCCGCTTCTCTCCCTGTGGGAAAGCGTGGGCTTCTCATAGTCAAGCTGTAGG 2125
 Db 2100 TACCTGTCCGCTTCTCTCCCTGTGGGAAAGCGTGGGCTTCTCATAGTCAAGCTGTAGG 2159
 QY 2126 TATCTCAAGTTGGGTAGGTGTTGCTGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTT 2185
 Db 2160 TATCTCAAGTTGGGTAGGTGTTGCTGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTT 2219
 QY 2186 CAGCCCGACGCGTGGGCTTATCCGGGTAACTATGCTCTTGAATCCAAACCGGTAAAGAC 2245
 Db 2220 CAGCCCGACGCGTGGGCTTATCCGGGTAACTATGCTCTTGAATCCAAACCGGTAAAGAC 2279
 QY 2246 GACTTATCGCCTAGTCGACGACGCACTGTAAACAGATTAGCAGAGCGAGTATGTAGC 2305
 Db 2280 GACTTATCGCCTAGTCGACGACGCACTGTAAACAGATTAGCAGAGCGAGTATGTAGC 2339
 QY 2306 GGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCTACACTAGAGGACAGTATTT 2365
 Db 2340 GGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCTACACTAGAGGACAGTATTT 2399
 QY 2366 GGTATCGCGCTCTGCTGAGACGATTACTTCGGAAAAAGAGTTGGAGTCTTTGATCC 2425
 Db 2400 GGTATCGCGCTCTGCTGAGACGATTACTTCGGAAAAAGAGTTGGAGTCTTTGATCC 2459
 QY 2426 GGGCAACAAACACCGCTGTAGGCGGTGTTTTTTTGTTCGAAAGCAGAGATTACGCGC 2485
 Db 2460 GGGCAACAAACACCGCTGTAGGCGGTGTTTTTTTGTTCGAAAGCAGAGATTACGCGC 2519
 QY 2486 AGAAAAAAGAGATCTCAAGAGATCTTTGATCTTTTTCAGGGGCTGACGCTCAGTGG 2545
 Db 2520 AGAAAAAAGAGATCTCAAGAGATCTTTGATCTTTTTCAGGGGCTGACGCTCAGTGG 2579
 QY 2546 AACGAAACTCAGCTTAAAGGATTTGGTCAATGAGATTAATCAAAAAGGATCTTACCCAG 2605
 Db 2580 AACGAAACTCAGCTTAAAGGATTTGGTCAATGAGATTAATCAAAAAGGATCTTACCCAG 2639
 QY 2606 ATCCTTTTAAATTAAGATTAAGATTTTAAATCAATCTAAAGTATATAGTAAACTTGG 2665
 Db 2640 ATCCTTTTAAATTAAGATTAAGATTTTAAATCAATCTAAAGTATATAGTAAACTTGG 2699
 QY 2666 TCTGACAGTTACCAATGCTTAATCAATGAGGACACTATCTGACGAGTGTCTATTTGGT 2725
 Db 2700 TCTGACAGTTACCAATGCTTAATCAATGAGGACACTATCTGACGAGTGTCTATTTGGT 2759
 QY 2726 TCAATCCATAGTTGGCTGATCTCCCGTGTGTAGATTAATCAATGAGGAGGCTTACCA 2785
 Db 2760 TCAATCCATAGTTGGCTGATCTCCCGTGTGTAGATTAATCAATGAGGAGGCTTACCA 2819
 QY 2786 TCTGGCCCAAGTGTGCAATGATACCGGAGACCAAGCTCAGCGGCTCCAGATTATCA 2845
 Db 2820 TCTGGCCCAAGTGTGCAATGATACCGGAGACCAAGCTCAGCGGCTCCAGATTATCA 2879
 QY 2846 GCAATTAACCAAGCAGCGGAGGAGGCGAGCGCAGAAAGTGTCTGCAACTTTATCCGCC 2905
 Db 2880 GCAATTAACCAAGCAGCGGAGGAGGCGAGCGCAGAAAGTGTCTGCAACTTTATCCGCC 2939
 QY 2906 TCCATCCAGTCTATTAATGTTGCGGGAAGCTAGAGTAATAGTCCCGCATTAATAGT 2965
 Db 2940 TCCATCCAGTCTATTAATGTTGCGGGAAGCTAGAGTAATAGTCCCGCATTAATAGT 2999
 QY 2966 TTGGGCAACGTTGTTGGCATTTGCTACAGGCAATCGTGTGTCAAGCTGCTTTGGTATG 3025
 Db 3000 TTGGGCAACGTTGTTGGCATTTGCTACAGGCAATCGTGTGTCAAGCTGCTTTGGTATG 3059

QY 3026 GCTTCATTACACTCCGTTCCCAACGATCAAGGCGGAGTTACATGATCCCCATGTTGTC 3085
 Db 3060 GCTTCATTACACTCCGTTCCCAACGATCAAGGCGGAGTTACATGATCCCCATGTTGTC 3119
 QY 3086 AAAAAAGCGGTAGCTCCCTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTTGGCCGACGTG 3145
 Db 3120 AAAAAAGCGGTAGCTCCCTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTTGGCCGACGTG 3179
 QY 3146 TTATCACTCATGTTATGACGACACTGATTAATCTCTTACTGTATGCCATCCGTAGA 3205
 Db 3180 TTATCACTCATGTTATGACGACACTGATTAATCTCTTACTGTATGCCATCCGTAGA 3239
 QY 3206 TGCCTTTCTGTGATCGTGTAGTACTCAACCAAGTCTTCTGAAGATAGTATGCGCGCA 3265
 Db 3240 TGCCTTTCTGTGATCGTGTAGTACTCAACCAAGTCTTCTGAAGATAGTATGCGCGCA 3299
 QY 3266 CCGAGTGTCTTGTCCCGCGGTCAATACGGGATTAATACCGGCGCACATAGGAGAACTTTA 3325
 Db 3300 CCGAGTGTCTTGTCCCGCGGTCAATACGGGATTAATACCGGCGCACATAGGAGAACTTTA 3359
 QY 3326 AAAGTGTCAATCAATGGAACAAAGTCTTCCGGGCGGAAAACTCAAGAGATCTTACCGCTG 3385
 Db 3360 AAAGTGTCAATCAATGGAACAAAGTCTTCCGGGCGGAAAACTCAAGAGATCTTACCGCTG 3419
 QY 3386 TTGAGATCAAGTTGATGATTAACCACTGTGTGCAACCAATGATCTTTCAGCATCTTTTACT 3445
 Db 3420 TTGAGATCAAGTTGATGATTAACCACTGTGTGCAACCAATGATCTTTCAGCATCTTTTACT 3479
 QY 3446 TTCAACGACGTTCTGGGGTGAACAAAAACAGAAAGGCAAAATGCGGCAAAAAAGGGAATA 3505
 Db 3480 TTCAACGACGTTCTGGGGTGAACAAAAACAGAAAGGCAAAATGCGGCAAAAAAGGGAATA 3539
 QY 3506 AGGCGCACACGAAATGTGAATACATACTCTTCTTTTCAATATTAATGAAGCAT 3565
 Db 3540 AGGCGCACACGAAATGTGAATACATACTCTTCTTTTCAATATTAATGAAGCAT 3599
 QY 3566 TATCAGGTTATTTGTCTCATAGCGGATACATATTTGAATATTTAGAAAAATTAACA 3625
 Db 3600 TATCAGGTTATTTGTCTCATAGCGGATACATATTTGAATATTTAGAAAAATTAACA 3659
 QY 3626 ATAGGGGTCCGCGCACATTTCCCGGAAAAAGTGCAC 3662
 Db 3660 ATAGGGGTCCGCGCACATTTCCCGGAAAAAGTGCAC 3696

RESULT 4
 US-09-503-222-6
 ; Sequence 6, Application US/09503222
 ; Patent No. 6265548
 ; GENERAL INFORMATION:
 ; APPLICANT: Pavlakis, George N.
 ; APPLICANT: Galanaris, George A.
 ; APPLICANT: Stauber, Roland H.
 ; APPLICANT: Vournakis, John N.
 ; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
 ; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/503,222
 ; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,538
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: -
LOCATION: 1..3699
OTHER INFORMATION: /note="DBSGFP"
US-09-503-222-6

Query Match 71.6%; Score 2620.6; DB 3; Length 3699;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;
QY 3 AAATGTAAAGGCTTAATTTTGTAAATTCGGTTAAATTTTGTAAATTCAGTCTAT
DB 3 AAATGTAAAGGCTTAATTTTGTAAATTCGGTTAAATTTTGTAAATTCAGTCTAT
QY 63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTAATAATCAAAAGATAGACCGAGA
DB 63 TTTTAAACCAATAGCGCGAAATCGGCAAAATCCCTTAATAATCAAAAGATAGACCGAGA
QY 123 TAGGGTGAAGTGTGTCGAGTTTGAACAAGATCACTAATTAAGAAAGATGAGTCTCA
DB 123 TAGGGTGAAGTGTGTCGAGTTTGAACAAGATCACTAATTAAGAAAGATGAGTCTCA
QY 123 TAGGGTGAAGTGTGTCGAGTTTGAACAAGATCACTAATTAAGAAAGATGAGTCTCA
DB 123 TAGGGTGAAGTGTGTCGAGTTTGAACAAGATCACTAATTAAGAAAGATGAGTCTCA
QY 183 ACGTCAAAAGGCGCAAAACCGTCTATACAGGGGATGCGCCCTACGTAACCATCAACCT
DB 183 ACGTCAAAAGGCGCAAAACCGTCTATACAGGGGATGCGCCCTACGTAACCATCAACCT
QY 183 ACGTCAAAAGGCGCAAAACCGTCTATACAGGGGATGCGCCCTACGTAACCATCAACCT
DB 183 ACGTCAAAAGGCGCAAAACCGTCTATACAGGGGATGCGCCCTACGTAACCATCAACCT
QY 243 AATCAAGTTTTTGGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCCCTAAAGGGAGCC
DB 243 AATCAAGTTTTTGGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCCCTAAAGGGAGCC
QY 243 AATCAAGTTTTTGGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCCCTAAAGGGAGCC
DB 243 AATCAAGTTTTTGGGGTTCGAGTGCCTTAAGCACTAAATCGGAACCCCTAAAGGGAGCC
QY 303 CCGAATTAAGAGCTTGAACGGGAAAGCGCGCAACGTGGCGAAGAAAGGAAAGAAAG
DB 303 CCGAATTAAGAGCTTGAACGGGAAAGCGCGCAACGTGGCGAAGAAAGGAAAGAAAG
QY 303 CCGAATTAAGAGCTTGAACGGGAAAGCGCGCAACGTGGCGAAGAAAGGAAAGAAAG
DB 303 CCGAATTAAGAGCTTGAACGGGAAAGCGCGCAACGTGGCGAAGAAAGGAAAGAAAG
QY 363 CGAAAGAGCGGGCGCTAGGGCGCTGCAAGTGTACGGGTCAAGCTGGCGCTAACCA
DB 363 CGAAAGAGCGGGCGCTAGGGCGCTGCAAGTGTACGGGTCAAGCTGGCGCTAACCA
QY 423 CACCCGCGCGCTTAATAGCGCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGC
DB 423 CACCCGCGCGCTTAATAGCGCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGC
QY 423 CACCCGCGCGCTTAATAGCGCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGC
DB 423 CACCCGCGCGCTTAATAGCGCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGC
QY 480 GCAACTTTGGGAAGGCGATCGGTGCGGCGCTTCCCTTAATACGCACTGGCGAAG
DB 480 GCAACTTTGGGAAGGCGATCGGTGCGGCGCTTCCCTTAATACGCACTGGCGAAG
QY 483 GCAACTTTGGGAAGGCGATCGGTGCGGCGCTTCCCTTAATACGCACTGGCGAAG
DB 483 GCAACTTTGGGAAGGCGATCGGTGCGGCGCTTCCCTTAATACGCACTGGCGAAG
QY 540 GGGGATGTCTGCAAGGCGATTAAGTGGGTAAACCGAGGGTTTCCCACTCAAGCGTT
DB 540 GGGGATGTCTGCAAGGCGATTAAGTGGGTAAACCGAGGGTTTCCCACTCAAGCGTT
QY 543 GGGGATGTCTGCAAGGCGATTAAGTGGGTAAACCGAGGGTTTCCCACTCAAGCGTT
DB 543 GGGGATGTCTGCAAGGCGATTAAGTGGGTAAACCGAGGGTTTCCCACTCAAGCGTT
QY 600 GTTAAACCAAGCGCGCTAGAGCGCGCTGCTTCAATCAAGTTTGAACCCGTTGAGAGCG
DB 600 GTTAAACCAAGCGCGCTAGAGCGCGCTGCTTCAATCAAGTTTGAACCCGTTGAGAGCG
QY 603 GTTAAACCAAGCGCGCTAGAGCGCGCTGCTTCAATCAAGTTTGAACCCGTTGAGAGCG
DB 603 GTTAAACCAAGCGCGCTAGAGCGCGCTGCTTCAATCAAGTTTGAACCCGTTGAGAGCG
QY 660 GGCAGACTCGCGGTGCAAAATGTGTTTACAGCGGTGAGAGAGATGAAGATGCTCGACA
DB 660 GGCAGACTCGCGGTGCAAAATGTGTTTACAGCGGTGAGAGAGATGAAGATGCTCGACA

DB 661 GCCCCCCCTGAGGCTGACGGTATTCGATTAACCTTGATGATCTTATTTGATAGTCAATC
QY 720 CGGTGCA---GAACACGAGCTAGATTAACCTTAAGAAAGTATCATATTTGAGAGTACG
DB 721 CATGCAATGTGATATCCGACAGCTGTAAACCTCAAGAGAGACATGTCTCTCTT
QY 777 TTAAGATTAATCATGCTGTAATAATTGACGAGTGGGATCTGTAATACACTCATATAGAGG
DB 781 TTCTGTGGATCTTTGAAAGGCGAATTTGTGACAGGTA-ATGTTGCTGTGTAATAA
QY 837 GAATTGGGTAACCGGCCCCCTCGAGGTGACAGGTATCGATTAAGCTTGAATATGAAATTC
DB 840 GGAAGAGGCGCATCGCAATTTGAGATATTTGTAATATGCTGTCTGTTGAAGCTTC
QY 897 CTGAGCGCGGGGATCACTAGTT---CTAGAGCGCGCGCACCGCGTGAAGCTCA
DB 900 CATCTTCATATGTGTCTAATTTTGAAGTTAATTTGATTCATTTCTTTGTTGCTG
QY 953 GCTTTTGTCCCTTAGAGAGGTTAATTAATGATCCATGCGTCAATTTTACGAGACTAT
DB 960 CCATGATGATACATTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT
QY 1013 CTTTCTAGGTTAATCTAGCTGATCAGATCATATGCTGGGCTTTTTCGGCTCAG
DB 1020 CATCTTTTAAATCAATACCTTTTAACTGATTTAATTAACAGGATACCTTCA
QY 1073 TATGCGCAAGCTGCGCTATCTGCGCATCGGGAGAGAAAGCCGCTTTTCCG
DB 1080 ACTTGACTTACAGAGTGTCTGATGTTCCGCTCATCTTTGAAATTAATGTTCTTCT
QY 1133 CGAGGTGAAGCGGATGAGAAAGTTTCCGAGATGATGATGATGATGATGATGATGATGAT
DB 1140 GTACATTAATCTTGGGAGAGCACTCTTGAAATGATGATGATGATGATGATGATGATGAT
QY 1193 GCGAAACGCACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT
DB 1200 ATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
QY 1253 TGAATGTTGTTCAAGGCGCACTGGGATACAGATGCTGCGGCTTTCCGGAACAAGT
DB 1260 GTAGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
QY 1313 CCGGATGTCAGCGGCAAGCCCATCAGCAACCGGAATAATCGGAGACAGCGGAAGT
DB 1320 CACCTTCTCACTGACGAAATTTGCGCCATTAATCATTACATCTTAATTAACAAGAA
QY 1373 CCGTCC-----GGTGCAGATTAATGACAGCGGTGCGCGCTGGGATAT
DB 1380 TTGGACAATCCAGTAAAGTTCTTCTCTTGTGCTACATTTCTTGGGATGAT
QY 1419 TACGTGACGAGAGCGGATATCTGCTGATGATGATGATGATGATGATGATGATGATGAT
DB 1440 TCTGCAAGCGCGGAGATCACTAGTTTGAAGCGCGCGCACCGGTGAGGCTCAGC
QY 1479 GAGTTACCGGCGG-----CGGCTTGGGATGATGATGATGATGATGATGATGATGAT
DB 1500 TTTTGTCTCTTAATGAGAGGTTAATTCGAGCTTGGGATGATGATGATGATGATGAT
QY 1526 CCTGTGTAATTTTATTCGCTCAAAATTCACAAACATGAGCGGAGACATTAAG
DB 1560 CTGTGTGTAATTTTATTCGCTCAAAATTCACAAACATGAGCGGAGACATTAAG
QY 1586 TGTAAAGCTTGGGATGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT
DB 1620 TGTAAAGCTTGGGATGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT
QY 1646 CCGGCTTTCAGAGGAGAACTGTGTGCAAGCTCATTAATGATGATGATGATGATGATGAT
DB 1680 CCGGCTTTCAGAGGAGAACTGTGTGCAAGCTCATTAATGATGATGATGATGATGATGAT
QY 1706 GGAAGAGCGGTTGCTATTTGGGCGCTCTCGCTTCTGCTCACTGACTGCTGCGC
DB 1765

Db 1740 GGGAGAGCGGTTTTCGATTATGGGCGCTCTTCCGCTCTCGCTCACTGACTCGCTGCGC 1799
Qy 1766 TCGGCTGTTCCGCTGCGGCGAGCGGTTATCACTCACTAAAGCGGTTAATAGGTTATCC 1825
Db 1800 TCGGCTGTTCCGCTGCGGCGAGCGGTTATCACTCACTAAAGCGGTTAATAGGTTATCC 1859
Qy 1826 ACAGATCAAGGAGATTAACGAGAAAGAAATGATGAGCAAAAGGCGAGAAAGGCGAG 1885
Db 1860 ACAGATCAAGGAGATTAACGAGAAAGAAATGATGAGCAAAAGGCGAGAAAGGCGAG 1919
Qy 1886 AACGTTAAAAAGCGCGCTGCTGCGGTTTTCATAGGCTTCCGCGCGCTGACAGCAT 1945
Db 1920 AACGTTAAAAAGCGCGCTGCTGCGGTTTTCATAGGCTTCCGCGCGCTGACAGCAT 1979
Qy 1946 CACAAAAATCGACCGTTCAAGTCAAGGTGGGAAACCCGACAGGACTTAAAGATACAG 2005
Db 1980 CACAAAAATCGACCGTTCAAGTCAAGGTGGGAAACCCGACAGGACTTAAAGATACAG 2039
Qy 2006 GCGTTTCCCGTGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCGCTGACCGGA 2065
Db 2040 GCGTTTCCCGTGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCGCTGACCGGA 2099
Qy 2066 TACCTGTCGCGCTTCTCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGTAAG 2125
Db 2100 TACCTGTCGCGCTTCTCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGTAAG 2159
Qy 2126 TATCTCAATTGGGTAGTGTGCTGCTTCCAAAGCTGGGCTGTGTCAAGAACCCCGCTT 2185
Db 2160 TATCTCAATTGGGTAGTGTGCTGCTTCCAAAGCTGGGCTGTGTCAAGAACCCCGCTT 2219
Qy 2186 CAGCCCGACGCTGCGCTTATCCGCTTAATGCTTGTGATCCGACCCGCTTAAGACAC 2245
Db 2220 CAGCCCGACGCTGCGCTTATCCGCTTAATGCTTGTGATCCGACCCGCTTAAGACAC 2279
Qy 2246 GACTTATCGCACTGGCAGACCACTGTAACAGATTAGCAGAGGATATGTAGGC 2305
Db 2280 GACTTATCGCACTGGCAGACCACTGTAACAGATTAGCAGAGGATATGTAGGC 2339
Qy 2306 GGTCTCAAGAGTTCTTGAAGTGTGGCTTAATACGCTCACTAGAAAGACAGTATTT 2365
Db 2340 GGTCTCAAGAGTTCTTGAAGTGTGGCTTAATACGCTCACTAGAAAGACAGTATTT 2399
Qy 2366 GGTATCTGCGCTGCTGTAAGCACTTCCGAAAAAGAGTTGGAGCTTTGATCC 2425
Db 2400 GGTATCTGCGCTGCTGTAAGCACTTCCGAAAAAGAGTTGGAGCTTTGATCC 2459
Qy 2426 GGCAGAACAAACACCGCTGTAGCGGTGTTTTTTTGTTCAGACAGCAATTAACGCGC 2485
Db 2460 GGCAGAACAAACACCGCTGTAGCGGTGTTTTTTTGTTCAGACAGCAATTAACGCGC 2519
Qy 2486 AGAAAAAAGAGATCTCAAGAAAGATCTTTTGTTCACGGGCTGACGCTCAAGTG 2545
Db 2520 AGAAAAAAGAGATCTCAAGAAAGATCTTTTGTTCACGGGCTGACGCTCAAGTG 2579
Qy 2546 AACGAAAACTCAGCTTAAGGATTTTGGTCAATGATTAATCAAAAGGATCTTACCTAG 2605
Db 2580 AACGAAAACTCAGCTTAAGGATTTTGGTCAATGATTAATCAAAAGGATCTTACCTAG 2639
Qy 2606 ATCTTTTAAATTAATTAATGAAGTTTAAATCAATCTAAAGTATATGATTAACCTGG 2665
Db 2640 ATCTTTTAAATTAATTAATGAAGTTTAAATCAATCTAAAGTATATGATTAACCTGG 2699
Qy 2666 TCTGACAGTTTACCAATGCTTATCAATGAGGCACTTATCTCAGGAGTGTGTATTTTCT 2725
Db 2700 TCTGACAGTTTACCAATGCTTATCAATGAGGCACTTATCTCAGGAGTGTGTATTTTCT 2759
Qy 2726 TCAATCAATGTTGCTGACTCCCGCTGTGTAGATTAATCTAGATTAAGGAGGCTTACCA 2785
Db 2760 TCAATCAATGTTGCTGACTCCCGCTGTGTAGATTAATCTAGATTAAGGAGGCTTACCA 2819
Qy 2786 TCTGGCCCAAGTGTGCAATGATACCGGAGAACCCAGCTCACCAGCTCCAGATTTATCA 2845
Db 2820 TCTGGCCCAAGTGTGCAATGATACCGGAGAACCCAGCTCACCAGCTCCAGATTTATCA 2879

Qy 2846 GCAATTAACCAAGCCAGCCGGAAGGCGGAGGCAAGAAAGTGGTCTTCAACTTTATCCGCC 2905
Db 2880 GCAATTAACCAAGCCAGCCGGAAGGCGGAGGCAAGAAAGTGGTCTTCAACTTTATCCGCC 2939
Qy 2906 TCAATCCAGTCTTATTAATTTGTTCGCGGAACTAGAGTATGATTTGCCAGTTAATAGT 2965
Db 2940 TCAATCCAGTCTTATTAATTTGTTCGCGGAACTAGAGTATGATTTGCCAGTTAATAGT 2999
Qy 2966 TTGGGCAACGTTGTCATTTGCTACAGGCACTGTGTGTACAGCTGCTGTTGGTATG 3025
Db 3000 TTGGGCAACGTTGTCATTTGCTACAGGCACTGTGTGTGTACAGCTGCTGTTGGTATG 3059
Qy 3026 GCTTCAATCACTCCGTTTCCCAACGATCAAGGCGAGTTATATGATCCCAATGTTGTC 3085
Db 3060 GCTTCAATCACTCCGTTTCCCAACGATCAAGGCGAGTTATATGATCCCAATGTTGTC 3119
Qy 3086 AAAAAAGCGGTAGCTCCTTCCGTCCTCCGATCGTTGTCAGAAAGTGTGGCCGACGTG 3145
Db 3120 AAAAAAGCGGTAGCTCCTTCCGTCCTCCGATCGTTGTCAGAAAGTGTGGCCGACGTG 3179
Qy 3146 TTATCACTCATGTTATGAGCAGCACTGCATTAATCTCTTACTGTATGCAATCCGTAGA 3205
Db 3180 TTATCACTCATGTTATGAGCAGCACTGCATTAATCTCTTACTGTATGCAATCCGTAGA 3239
Qy 3206 TGTCTTTCTGTGACTGTGTGATCTCAACCAAGTCAATTTGGAATATGTTATGCGCGA 3265
Db 3240 TGTCTTTCTGTGACTGTGTGATCTCAACCAAGTCAATTTGGAATATGTTATGCGCGA 3299
Qy 3266 CCGAGTTGCTCTTCCGCGGCGTCAATACGGGATTAATCCGCGCAATAGAGAACTTTA 3325
Db 3300 CCGAGTTGCTCTTCCGCGGCGTCAATACGGGATTAATCCGCGCAATAGAGAACTTTA 3359
Qy 3326 AAAGTGTCAATCATTTGAAAAAGTTCTTCCGCGGCGAAAACTTCAAGATCTTACCGCTG 3385
Db 3360 AAAGTGTCAATCATTTGAAAAAGTTCTTCCGCGGCGAAAACTTCAAGATCTTACCGCTG 3419
Qy 3386 TTGAGATTCAGATTGATGATTAACCACTGTGTGACCCCACTGATTTTGACATCTTTTACT 3445
Db 3420 TTGAGATTCAGATTGATGATTAACCACTGTGTGACCCCACTGATTTTGACATCTTTTACT 3479
Qy 3446 TTCAACAGCGTTTCTGCGTGTGCAAAAAAGGAGGCAAAATGCGGCAAAAGGAGATA 3505
Db 3480 TTCAACAGCGTTTCTGCGTGTGCAAAAAAGGAGGCAAAATGCGGCAAAAGGAGATA 3539
Qy 3506 AGGCGCACAGGAAATGTAATCACTATCTCTTCTTTTCAATATTAATTAAGGATTT 3565
Db 3540 AGGCGCACAGGAAATGTAATCACTATCTCTTCTTTTCAATATTAATTAAGGATTT 3599
Qy 3566 TATCAGGCTTATTTGTCTCATGAGCGGATTAATATTTGAATGATTTGAAGAAATTAACA 3625
Db 3600 TATCAGGCTTATTTGTCTCATGAGCGGATTAATATTTGAATGATTTGAAGAAATTAACA 3659
Qy 3626 ATAGGGGTCCGCGCACTTTCCCGGAAAGTGCAC 3662
Db 3660 ATAGGGGTCCGCGCACTTTCCCGGAAAGTGCAC 3696

RESULT 5
US-09-890-16
Sequence 16 Application US/09890848
Patent No. 6573053
GENERAL INFORMATION:
APPLICANT: Fitch, Greg
APPLICANT: Odeira, Rajesh Muru
TITLE OF INVENTION: Analysis Method
FILE REFERENCE: P93903
CURRENT APPLICATION NUMBER: US/09/890,848
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: PCT/GB00/00346
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: BP 99300873.9
PRIOR FILING DATE: 1999-02-05

QY 1980 ACCGACAGACTATAAAGATACAGAGGTTTCCCTGAGAGTCCCTGCGCTCTC 2039
DB ACCGACAGACTATAAAGATACAGAGGTTTCCCTGAGAGTCCCTGCGCTCTC 3068
QY 2040 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAACGTTG 2099
DB CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAACGTTG 3128
QY 2100 CGCTTCTCATAGCTACAGCTGTAGATATCTCAGTTCCGTTGAGTGTGCTTCCCAAGC 2159
DB CGCTTCTCATAGCTACAGCTGTAGATATCTCAGTTCCGTTGAGTGTGCTTCCCAAGC 3188
QY 2160 TGGGCTGTGTCAGAGAACCCCGCTTACGCGGACCGGCTGCGCTTATCCGCTTACTAC 2219
DB TGGGCTGTGTCAGAGAACCCCGCTTACGCGGACCGGCTGCGCTTATCCGCTTACTAC 3248
QY 2220 GTCTTGAATCCAAACCCGTTAAGACACGACTTATCCGCACTGGCAGACGCACTGGTAACA 2279
DB GTCTTGAATCCAAACCCGTTAAGACACGACTTATCCGCACTGGCAGACGCACTGGTAACA 3308
QY 2280 GGATTTAGCAGACGAGATATGATAGCGGCTGCTACAGAGTTCTTGAAGTGTGCTTAACT 2339
DB GGATTTAGCAGACGAGATATGATAGCGGCTGCTACAGAGTTCTTGAAGTGTGCTTAACT 3368
QY 2340 ACGGCTACACTAGAAAGACAGATTTGGTATCTGGGCTCTGTGAGGCACTTACTTCG 2399
DB ACGGCTACACTAGAAAGACAGATTTGGTATCTGGGCTCTGTGAGGCACTTACTTCG 3428
QY 2400 GAAAAAGAGTTGTAGCTCTTGATCCGCAAAACCAACGCGTGTAGCGGTGTTT 2459
DB GAAAAAGAGTTGTAGCTCTTGATCCGCAAAACCAACGCGTGTAGCGGTGTTT 3488
QY 2460 TTGTTTGAACAGACAGATTAACGCGCAGAAAAAAGATCTCAGAGAGATCTTGTATCT 2519
DB TTGTTTGAACAGACAGATTAACGCGCAGAAAAAAGATCTCAGAGAGATCTTGTATCT 3548
QY 2520 TTTCTAGCGGGCTGACGCTCAGTGAAGAAACTCAGTTAAGGATTTTGGTCAAGA 2579
DB TTTCTAGCGGGCTGACGCTCAGTGAAGAAACTCAGTTAAGGATTTTGGTCAAGA 3608
QY 2580 GATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAA 2639
DB GATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAA 3668
QY 2640 TCTAAGTATATATGATTAACCTTGTCTGACAGTTACCAATGCTTAATCAGTAGCAGC 2699
DB TCTAAGTATATATGATTAACCTTGTCTGACAGTTACCAATGCTTAATCAGTAGCAGC 3728
QY 2700 CTATCTCAGGATCTGTCTATTTGGTTATCTCATAGTTGCTGACTCCCGCTCGTTAGA 2759
DB CTATCTCAGGATCTGTCTATTTGGTTATCTCATAGTTGCTGACTCCCGCTCGTTAGA 3788
QY 2760 TAACTACGATACGGAAGGCTTACCATCTGCGCCAGAGTGTGCAATGATACCGGAGACC 2819
DB TAACTACGATACGGAAGGCTTACCATCTGCGCCAGAGTGTGCAATGATACCGGAGACC 3848
QY 2820 CAGCTCAGCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGGAGCGCA 2879
DB CAGCTCAGCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGGAGCGCA 3908
QY 2880 GAACTGTCTCCGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGTA 2939
DB GAACTGTCTCCGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGTA 3968
QY 2940 GAGTAAGTATGTCGCAAGTTATGTTGCGCAAGTTGTCATGCTCAAGGATCG 2999
DB GAGTAAGTATGTCGCAAGTTATGTTGCGCAAGTTGTCATGCTCAAGGATCG 4028
QY 3000 TGGTGTACGCTGCTGTTGTGATGCTTCACTCACTCGGTTCCCAAGATCAAGGC 3059
DB TGGTGTACGCTGCTGTTGTGATGCTTCACTCACTCGGTTCCCAAGATCAAGGC 4088

QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGGCTTAGAGCTTCCGTCCTCCGATG 3119
DB GAGTTACATGATCCCCCATGTTGTGCAAAAAAGGCTTAGAGCTTCCGTCCTCCGATG 4148
QY 3120 TTGTGAGAGTATGTTGCGCCAGTGTATCACTCATATGAGTATGAGCAGTCAATAT 3179
DB TTGTGAGAGTATGTTGCGCCAGTGTATCACTCATATGAGTATGAGCAGTCAATAT 4208
QY 3180 CTCTTACTGTATGCTCATCCGTAAGATGCTTTCTGTGACTGTGTAGTATCAACCAAGT 3239
DB CTCTTACTGTATGCTCATCCGTAAGATGCTTTCTGTGACTGTGTAGTATCAACCAAGT 4268
QY 3240 CATCTGAGATATGATATGCGGCGACCGAGTGTCTTGGCCGCGCAATACGAGTA 3299
DB CATCTGAGATATGATATGCGGCGACCGAGTGTCTTGGCCGCGCAATACGAGTA 4328
QY 3300 ATACCGCCCATATAGCAGAACTTTAAAGTCTCATCATTTGAAAGCGTTCTGCGGC 3359
DB ATACCGCCCATATAGCAGAACTTTAAAGTCTCATCATTTGAAAGCGTTCTTCCGCGC 4388
QY 3360 GAAAACTCTAAGATCTTACCGCTGTGAGATCAAGTTCGATTAACCACTCGTGAC 3419
DB GAAAACTCTAAGATCTTACCGCTGTGAGATCAAGTTCGATTAACCACTCGTGAC 4448
QY 3420 CCAACTGATCTTACGATCTTTTACTTTTCAACAGGTTTCCGCTGAGCAAAACAGGA 3479
DB CCAACTGATCTTACGATCTTTTACTTTTCAACAGGTTTCCGCTGAGCAAAACAGGA 4508
QY 3480 GGCAGAAATGCGGCAAAAGGAAATAGGCGGACACGGAATGTTGAATCTCATCT 3539
DB GGCAGAAATGCGGCAAAAGGAAATAGGCGGACACGGAATGTTGAATCTCATCT 4568
QY 3540 TCTTTTCAATATTAATTAAGCAATTAATCAGGTTATGTTCTCATGAGCGGATATAT 3599
DB TCTTTTCAATATTAATTAAGCAATTAATCAGGTTATGTTCTCATGAGCGGATATAT 4628
QY 3600 TTGAATGATTTAATAAATAAACAATAGGCGTTCCGCGCAATTTCCCGAAGAGTC 3659
DB TTGAATGATTTAATAAATAAACAATAGGCGTTCCGCGCAATTTCCCGAAGAGTC 4688
QY 3660 CAC 3662
DB 4689 CAC 4691

RESULT 6
US-09-402-266B-21
Sequence 21, Application US/09402266B
Patent No. 6537767
GENERAL INFORMATION:
APPLICANT: HINNEN, Albert
APPLICANT: HEGEMANN, Johannes
APPLICANT: MUNDER, Thomas
APPLICANT: SCHUSTER, Tilmer
APPLICANT: FELDMAN, Horst
APPLICANT: KRAMER, Wilfried
APPLICANT: ZIMMERMANN, Friedrich
APPLICANT: ENTIAN, Karl-Dieter
TITLE OF INVENTION: Process for Screening Antimycotically Active Substances
FILE REFERENCE: 38005-0094
CURRENT APPLICATION NUMBER: US/09/402,266B
CURRENT FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: PCT/EP98/01904
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: DE 19713572.2
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.0
SEQ ID NO 21
LENGTH: 3956
TYPE: DNA
ORGANISM: Unknown
FEATURE:

NAME/KEY: misc feature
LOCATION: (1..7)
OTHER INFORMATION: Synthetic sequence
us-09-402-266B-21

Query Match
Best Local Similarity 66.9%; Score 2448.8; DB 4; Length 3956;
Matches 3136; Conservative 0; Mismatches 522; Indels 294; Gaps 8;

QY 5 ATTGTAAGCGTTAAATTTTGTAAATTCGGTTAAATTTTGTAAATTCAGTCAATTT 64
DB 1 ATTGTAAGCGTTAAATTTTGTAAATTCGGTTAAATTTTGTAAATTCAGTCAATTT 64
QY 65 TTAAACCAATAGCCGGAATCGGCAAAATCCCTTAATATCAAAAGAAATAGACCGAT 60
DB 61 TTAAACCAATAGCCGGAATCGGCAAAATCCCTTAATATCAAAAGAAATAGACCGAT 60
QY 125 GGGTGAAGTGTGTTCCAGTTTGAACAAGTCCATATTAAGAAAGTGAATCCCAAC 184
DB 121 GGGTGAAGTGTGTTCCAGTTTGAACAAGTCCATATTAAGAAAGTGAATCCCAAC 184
QY 185 GTCAAAAGGCGAAAACCGTCTATCAAGGCGATGCGCCATCACTGAAACATCAACCTTA 244
DB 181 GTCAAAAGGCGAAAACCGTCTATCAAGGCGATGCGCCATCACTGAAACATCAACCTTA 244
QY 245 TCAAGTTTTTGGGGTGAAGTGTGCGTAAAGCACTAAATCGGAACCTTAAGGAGGCC 304
DB 241 TCAAGTTTTTGGGGTGAAGTGTGCGTAAAGCACTAAATCGGAACCTTAAGGAGGCC 304
QY 305 CGATTTAGAGCTTGAAGGCGAAAGCCGCGAAGTGTGCGTAAAGCACTAAATCGGAAC 300
DB 301 CGATTTAGAGCTTGAAGGCGAAAGCCGCGAAGTGTGCGTAAAGCACTAAATCGGAAC 300
QY 365 AAAAGAGCGGCGCTGAGGCGCTGCAAGTGTGCGTAAAGCACTAAATCGGAACCTTA 360
DB 361 AAAAGAGCGGCGCTGAGGCGCTGCAAGTGTGCGTAAAGCACTAAATCGGAACCTTA 360
QY 425 CCGCGCGCTTAAATGCGCGCTCAAGGCGCGCTCCATTCGCAATTCAGGCTGCGAAC 484
DB 421 CCGCGCGCTTAAATGCGCGCTCAAGGCGCGCTCCATTCGCAATTCAGGCTGCGAAC 484
QY 485 TGTGGGAAGGCGATGCTGCGCGCTCTCGCTTAACGCACTGCGCAAGGCGGGA 544
DB 481 TGTGGGAAGGCGATGCTGCGCGCTCTCGCTTAACGCACTGCGCAAGGCGGGA 544
QY 545 TGTGCTCAAGCGATTAAGTGTGGAAGCGCAAGGTTTTTCCAGTCAAGCGTTGTA 604
DB 541 TGTGCTCAAGCGATTAAGTGTGGAAGCGCAAGGTTTTTCCAGTCAAGCGTTGTA 604
QY 605 ACAGCGCGAGTGAAGCGCGCTCGTTCACTGCTTTTGAACCGCGAGGAGCGGCG 664
DB 601 ACAGCGCGAGTGAAGCGCGCTCGTTCACTGCTTTTGAACCGCGAGGAGCGGCG 664
QY 665 ACTGCGGCTGCAATGTGTTTACAGCGTATGAGC----- 701
DB 661 GGTGGCGCTTGAAGTGTGATTCGCGCAATGAGGCGCAACCGCATGATCGCAAGT 720
QY 702 ----- 701
DB 721 GCACAAACATACCTTAATACTACTAGTAATACCTATTTCTTAAGATTTTGGAC 780
QY 702 -----AGATGAAGATGCTGACACGCTGAGAAACCGACGATAGATT 743
DB 781 GAATTTTGTATTTTGTAGAGCTTTTACCACTTTTGTCTCACACTCGCTTACATC 840
QY 744 AACCTTGAAGAAATATCATATTTGAGCGTAAAGATTAATCATATGCGT----- 794
DB 841 AACCAATTAAGCGCTTAATCTTAAAGCGATCAACCAATTTTCTGGGCTCAGTCCACC 900
QY 795 -----AAATTAAGCGATGAGATCTGTAATTAAGCACTCAATAGGCGAATT--- 841
DB 901 AGTAACATTAATTAAGCTTTTGGGCGCTCTCTTGTCTTCAACCAAGTCAAGAAATCGA 960

QY 842 ----- 841
DB 961 GTTCAATCCAAAGTTCACCTGTCCACCTGCTTGTGATCAACAAAGGAAATTAACGA 1020
QY 842 -----GGTACCGGCGCCCCCTCGAGGTGACGCTTATGATTAAGCTTGAATTC 897
DB 1021 ATAGAGTTTCTGTGAAGCTGACGAGTGAATGATGATGATGATGATGATGATGATG 1080
QY 898 TGCAGCCGCGGAGATCCATGCTTGAAGCGCGCCACCGCGGTGAGTCCA----- 952
DB 1081 TTAAATACCTGCAACCGAGAACTTTGATATTTCTGCAAGTCAATCTCAATGAG 1140
QY 953 --GCTTTGTTCCCTTATGAGAGGTTAATTAATCCATGCTCAATTTTAAGCAGACT 1010
DB 1141 TTGACGATATCAATGCGGTATCATATGACAGGCAAAACATCTCTTAGTGTAT 1200
QY 1011 ATCTTTC----- 1017
DB 1201 ACGAAACGCGCAACCAAGTATTTGAGAGTGCCTGAACATATTTTATGCTTTACAG 1260
QY 1018 ---TAGGTTAATCTGATGATCAAGATCAATATGCTGAGCTTTTCCGCTCAGTC 1074
DB 1261 ACTTGAATTTTCTTGTCAATTAACCGGCTCAATTTGTTCTTTTATTTGAGGCA 1320
QY 1075 ATGCGCGAAGCTGCGCTATCTG-----GCAATCGGCGAGAAAGAA 1115
DB 1321 AATATCCCAAGTCAAGATCGAATCTGAGCAATCTGAGCAATTTAATACAGACATCTCA 1440
QY 1116 GCGCGTCTTTTCCGCGAGGTTGAAGCGGATGAGAAAGATTTGCGAGATGACCTGC 1175
DB 1381 GCGCGAATTTTCAACCAATGAGCAAGCAATCTGAGCAATTTAATACAGACATCTCA 1440
QY 1176 TGTGCTATTTGAGTGAAG-----CGAAACGCAAGTTTACATGATGATTTGGG 1224
DB 1441 AGCTGCTTGTGTGCTTAATCAAGTATCACTGCTCAATATCAATGCTCAATGCTCC 1500
QY 1225 -----AAGTGTGCGCAATGACGCTTTTAAAGGTAAGTGTGCTGAGCG 1270
DB 1501 CTCTTGGCCCTCTCTCTTTTCTTTTTCACGCAATTAATTTCTTAATGGCAAAAAAGAA 1560
QY 1271 CACTTGGGAATCAAGTTCCTGCGGCTTTTCCGAGCAAGTTCGAGTGTGAGCGCCGA 1330
DB 1561 AAGTTCGATCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1331 GCGCATCAAGCAACCGCAATTAACCGGCAAGCGGAACTGCGGCTGAGTGTGAT 1390
DB 1621 CACTTGTGCTATGCTGCTGATTAATCTGAAAGTACATATTAATGATGATGATGAT 1680
QY 1391 TAATGACAGCGGTGCGCGCTGAGATTTTACGTCAGGAGGAGGATATCTGCTGAT 1450
DB 1681 AATATCGGCGCTGAGAAATTCATATCAAGCTTATGATACGTCAGGAGGAGG 1740
QY 1451 GCGCGAAGAAATGACATGATTAACCGGTAAGTTAACCGGCGGCGGCTTGGGTAATCA 1510
DB 1741 GCGCGGAGCAAGCTTTTGTCTTTAGTGAAGGTTAATTTGCGGCTTGGGTAATCA 1800
QY 1511 TGTGCTATGCTGTTTCTGTTGTAATTTTATTCGCTCAATTTTCAACCAACATTCGA 1860
DB 1801 TGTGCTATGCTGTTTCTGTTGTAATTTTATTCGCTCAATTTTCAACCAACATTCGA 1920
QY 1571 GCGGGAAGCAATTAAGTGAAGCTTGGGATCTTAATGAGGCTTAATCAATTAATTT 1630
DB 1861 GCGGGAAGCAATTAAGTGAAGCTTGGGATCTTAATGAGGCTTAATCAATTAATTT 1920
QY 1631 GCGTGGCTCACTGCTGCTTTTCAAGTGGGAAACCTGTCGCGCAAGCTGATTAAGA 1690
DB 1921 GCGTGGCTCACTGCTGCTTTTCAAGTGGGAAACCTGTCGCGCAAGCTGATTAAGA 1750
QY 1691 ATGCGCAACGCGGAGGAGAGGCGGTTTGCATTTGAGGCGCTTTCCGCTTCCGCTC 1750
DB 1981 ATGCGCAACGCGGAGGAGAGGCGGTTTGCATTTGAGGCGCTTTCCGCTTCCGCTC 2040
QY 1751 ACTGATCGCTGCGCTGATGCTTGGGCGAGGCGGATATGATGATCACTCAAGGCG 1810

Db 2041 ACTGCTGCTGCGCTCGGCTGCTGCTGCGGCGAGCGGATATGAGCTCACTCAAAAGCG 2100
|||
Qy 1811 GTAATACGGTTATCCAGAAATCAGGGGATAAACGAGAAAGAAACATGTGAGCAAAAGC 1870
|||
Db 2101 GTAATACGGTTATCCAGAAATCAGGGGATAAACGAGAAAGAAACATGTGAGCAAAAGC 2160
|||
Qy 1871 CAGCAAAAGGCCAGAAACCGTTAAAAAGCCGCTGCTGCTGCTTTTTCATAGCTCCGC 1930
|||
Db 2161 CAGCAAAAGGCCAGAAACCGTTAAAAAGCCGCTGCTGCTGCTTTTTCATAGCTCCGC 2220
|||
Qy 1931 CCCCCGAGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGCGAAACCCGACAGA 1990
|||
Db 2221 CCCCCGAGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGCGAAACCCGACAGA 2280
|||
Qy 1991 CTATTAAGATACAGGCGCTTTCCCGTGGAAAGCTCCCTGTCGCTCTCTGTTCCGAC 2050
|||
Db 2281 CTATTAAGATACAGGCGCTTTCCCGTGGAAAGCTCCCTGTCGCTCTCTGTTCCGAC 2340
|||
Qy 2051 CTGCGCTTACCGGATACCTGCTCGGCTTTCTCCCTTGGGAAAGGTGCGCTTTCAT 2110
|||
Db 2341 CTGCGCTTACCGGATACCTGCTCGGCTTTCTCCCTTGGGAAAGGTGCGCTTTCAT 2400
|||
Qy 2111 AGCTCAGCTGAGGATCTCAGTTGCGGTAGGTGCTTCCGCTCCAGCTGGGCTGTGTG 2170
|||
Db 2401 AGCTCAGCTGAGGATCTCAGTTGCGGTAGGTGCTTCCGCTCCAGCTGGGCTGTGTG 2460
|||
Qy 2171 CAGCAACCCCGCTTACGCGCGAGCCGCTGCGCTTATCCGATATCTATGTTAGTCC 2230
|||
Db 2461 CAGCAACCCCGCTTACGCGCGAGCCGCTGCGCTTATCCGATATCTATGTTAGTCC 2520
|||
Qy 2231 AACCCGTTAAGACACGACTTATCCGCTAGCGAGAGCCACTGTGTAACAGATTAGCGA 2290
|||
Db 2521 AACCCGTTAAGACACGACTTATCCGCTAGCGAGAGCCACTGTGTAACAGATTAGCGA 2580
|||
Qy 2291 GCGAGTATGAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCCCTAACTACGCGCTACAT 2350
|||
Db 2581 GCGAGTATGAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCCCTAACTACGCGCTACAT 2640
|||
Qy 2351 AGAAGACAGTATTTGATCTGCGCTGCTGCGAGCCAGTTACCTTCGAAAAAGAGTT 2410
|||
Db 2641 AGAAGACAGTATTTGATCTGCGCTGCTGCGAGCCAGTTACCTTCGAAAAAGAGTT 2700
|||
Qy 2411 GGTAGCTTTGATCCGGCAAAACCAACCGCTGCTGCTGCTTTTGTGTTGCAAG 2470
|||
Db 2701 GGTAGCTTTGATCCGGCAAAACCAACCGCTGCTGCTGCTTTTGTGTTGCAAG 2760
|||
Qy 2471 CAGCAGATTAACGCGAGAAAAAGGATCTCAAGAGATCCCTTGAATCTTTTCTACGGGG 2530
|||
Db 2761 CAGCAGATTAACGCGAGAAAAAGGATCTCAAGAGATCCCTTGAATCTTTTCTACGGGG 2820
|||
Qy 2531 TCTGACGCTCAGTGAGCAAAACCTCAGCTTAAGGAGTTTGTGTCATGAGATTATCAAA 2590
|||
Db 2821 TCTGACGCTCAGTGAGCAAAACCTCAGCTTAAGGAGTTTGTGTCATGAGATTATCAAA 2880
|||
Qy 2591 AGGATCTTACCTAGATCTTTTAAATTAAATGAAGTTTAAATCAATCTAAAGTATA 2650
|||
Db 2881 AGGATCTTACCTAGATCTTTTAAATTAAATGAAGTTTAAATCAATCTAAAGTATA 2940
|||
Qy 2651 TATAGTAACTTGGTCTGACAGTAAACCAATGCTTAAATCAGTGAAGGCACTATCTCAGCG 2710
|||
Db 2941 TATAGTAACTTGGTCTGACAGTAAACCAATGCTTAAATCAGTGAAGGCACTATCTCAGCG 3000
|||
Qy 2711 ATCTGTCTATTTCGTTCAATCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2770
|||
Db 3001 ATCTGTCTATTTCGTTCAATCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
|||
Qy 2771 CCGGAGGGCTTACCATCTGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2830
|||
Db 3061 CCGGAGGGCTTACCATCTGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
|||
Qy 2831 GCTCAGATTATCAGCAATAAACAGCAGCGGAGGCGGAGCGGAGGAGTGTGCT 2890
|||

Db 3121 GCTCAGATTATCAGCAATAAACAGCAGCGGAGGCGGAGCGGAGGAGTGTGCT 3180
|||
Qy 2891 GCAACTTATCCGCTCCATCCATCACTTATTAATTGTTGCCGGAAAGCTAGAGTAAGT 2950
|||
Db 3181 GCAACTTATCCGCTCCATCCATCACTTATTAATTGTTGCCGGAAAGCTAGAGTAAGT 3240
|||
Qy 2951 TCGCAGTATTAATGTTGCGCAAGCTGTTGCCCAATGCTTACAGGCACTGCTGCTGCTGCTGCT 3010
|||
Db 3241 TCGCAGTATTAATGTTGCGCAAGCTGTTGCCCAATGCTTACAGGCACTGCTGCTGCTGCTGCT 3300
|||
Qy 3011 TCGTGTGTTGATGCTTCAATCAGCTCCGCTTCCCAACGATCAAGCGAGTATCATGA 3070
|||
Db 3301 TCGTGTGTTGATGCTTCAATCAGCTCCGCTTCCCAACGATCAAGCGAGTATCATGA 3360
|||
Qy 3071 TCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCCGCTCCGATCGTTGTGAGAGT 3130
|||
Db 3361 TCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCCGCTCCGATCGTTGTGAGAGT 3420
|||
Qy 3131 AAGTTGCGCGAGTTATCACTGATGTTATGCGACAGCTGCAATTAATCTTACTGTC 3190
|||
Db 3421 AAGTTGCGCGAGTTATCACTGATGTTATGCGACAGCTGCAATTAATCTTACTGTC 3480
|||
Qy 3191 ATGCCATCCGTAAGATGCTTTCTGTGACTGTGATCAACCAAGTCAATCTGAGAA 3250
|||
Db 3481 ATGCCATCCGTAAGATGCTTTCTGTGACTGTGATCAACCAAGTCAATCTGAGAA 3540
|||
Qy 3251 TAGTGTATGCGGCAACGAGTGTCTTTGCCGCGCTCAATACGGGATTAATCCGCGCA 3310
|||
Db 3541 TAGTGTATGCGGCAACGAGTGTCTTTGCCGCGCTCAATACGGGATTAATCCGCGCA 3600
|||
Qy 3311 CATAGCAAACTTTAAAGTCTCATATGGAAGAGTTTGTGGGCGGAAACTCTCA 3370
|||
Db 3601 CATAGCAAACTTTAAAGTCTCATATGGAAGAGTTTGTGGGCGGAAACTCTCA 3660
|||
Qy 3371 AGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGGCAACCACTGATCT 3430
|||
Db 3661 AGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGGCAACCACTGATCT 3720
|||
Qy 3431 TCAGCATCTTTTACTTTCACCAAGCTTTCTGGGTGAGCAAAACAGAAAGCAAAATGCC 3490
|||
Db 3721 TCAGCATCTTTTACTTTCACCAAGCTTTCTGGGTGAGCAAAACAGAAAGCAAAATGCC 3780
|||
Qy 3491 GCAAAAAAGGAAATTAAGGCGCAACGGAATGTTGAATCTCAATCTTCTTTTCA 3550
|||
Db 3781 GCAAAAAAGGAAATTAAGGCGCAACGGAATGTTGAATCTCAATCTTCTTTTCA 3840
|||
Qy 3551 TATATTGAGCATTTATCAGGAGTTATGTTCTCATGAGCGGATCATATTGGAATGAT 3610
|||
Db 3841 TATATTGAGCATTTATCAGGAGTTATGTTCTCATGAGCGGATCATATTGGAATGAT 3900
|||
Qy 3611 TAGAAAAATAAACAAATAGGGGTTCCGCGCACTTTCCCGAAAAAGTCCAC 3662
|||
Db 3901 TAGAAAAATAAACAAATAGGGGTTCCGCGCACTTTCCCGAAAAAGTCCAC 3952
|||

RESULT 7
US-09-402-266B-20
; Sequence 20, Application US/09402266B
; Patent No. 6537767
; GENERAL INFORMATION:
; APPLICANT: HINNEN, Albert
; APPLICANT: HESEMANN, Johannes
; APPLICANT: MUNDER, Thomas
; APPLICANT: SCHUSTER, Tilmer
; APPLICANT: FELDLMANN, Horst
; APPLICANT: KRAMER, Wilfried
; APPLICANT: ZIMMERMANN, Friedrich
; APPLICANT: ENTIAN, Karl-Dieter
; TITLE OF INVENTION: Process for Screening Anticancerically Active Substances
; FILE REFERENCE: 38005-0094
; CURRENT APPLICATION NUMBER: US/09/402,266B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: PCT/EP98/01904

/ PRIOR FILING DATE: 1998-10-08
 / PRIOR APPLICATION NUMBER: DE 19713572.2
 / PRIOR FILING DATE: 1997-04-02
 / NUMBER OF SEQ ID NOS: 38
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 20
 / LENGTH: 4102
 / TYPE: DNA
 / ORGANISM: Unknown
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (..)
 / OTHER INFORMATION: Synthetic sequence
 / US-09-402-266B-20

Query Match 62.9%; Score 2305.2; DB 4; Length 4102;
 Best Local Similarity 76.9%; Pred. No. 0;
 Matches 3150; Conservative 0; Mismatches 508; Indels 440; Gaps 10;

```

QY 5 ATTGTAAGCGTTAATATTTTGTAAATTCGGCTTAATTTTGTAAATCACTCATTT
Db 1 ATTGTAAGCGTTAATATTTTGTAAATTCGGCTTAATTTTGTAAATCACTCATTT
QY 65 TTTAACCAATAGGCGCAAAATCGCAAAATCCCTTAATCAAAAGATAGCCGAGATA
Db 61 TTTAACCAATAGGCGCAAAATCGCAAAATCCCTTAATCAAAAGATAGCCGAGATA
QY 125 GGGTTAGTGTGTTCCAGTTTGAACAAGTCCACTATTTAAAGACGTGAC
Db 121 GGGTTAGTGTGTTCCAGTTTGAACAAGTCCACTATTTAAAGACGTGAC
QY 185 GTCAAAAGGCGCAAAATCGCTTATCAGGCGATGCGCCACTAGTAAACCATCCTTA
Db 181 GTCAAAAGGCGCAAAATCGCTTATCAGGCGATGCGCCACTAGTAAACCATCCTTA
QY 245 TCAAGTTTTTGGGGTGAAGTGCCTTAAGCACTAATTCGAAACCTTAAGGAGCC
Db 241 TCAAGTTTTTGGGGTGAAGTGCCTTAAGCACTAATTCGAAACCTTAAGGAGCC
QY 305 CGATTTAGAGCTTGAACGGGAAACCGCGCAACGTGCGAAGAAAGGAAAGGAGCC
Db 301 CGATTTAGAGCTTGAACGGGAAACCGCGCAACGTGCGAAGAAAGGAAAGGAGCC
QY 365 AAAGAGGCGGCGCTAAGGCGCTGCAAGTGAAGGTCACGCTGCGCGTAACACACACA
Db 361 AAAGAGGCGGCGCTAAGGCGCTGCAAGTGAAGGTCACGCTGCGCGTAACACACACA
QY 425 CCGCGCGCTTAATAGCGCGCTAAGGCGCGTCCATTGCGCATTTCAAGGCTGCGCAAC
Db 421 CCGCGCGCTTAATAGCGCGCTAAGGCGCGTCCATTGCGCATTTCAAGGCTGCGCAAC
QY 485 TGTGGGAAGGCGCATCGGTGCGGCTCTTTCGTAATTACGCGCACTGCGAAGGAGGA
Db 481 TGTGGGAAGGCGCATCGGTGCGGCTCTTTCGTAATTACGCGCACTGCGAAGGAGGA
QY 545 TGTGTGCAAGCGATTAAAGTGGGTAAGCCAGGGTTTTTCCAGTCAAGAGTTGTAA
Db 541 TGTGTGCAAGCGATTAAAGTGGGTAAGCCAGGGTTTTTCCAGTCAAGAGTTGTAA
QY 605 ACGAGCGCGAGTGAAGCGCGCTGTTTCAATTCAGTTTGAACCGGTGAGAGAGCGGAG
Db 601 ACGAGCGCGAGTGAAGCGCGCTGTTTCAATTCAGTTTGAACCGGTGAGAGAGCGGAG
QY 665 ACTCGCGGTGAATGTGTTTACAGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 661 GGTGGCGCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
QY 711 ----- 710
Db 721 TCAATTCATCATTTTTTTTTTATCTTTTTTGAATTCGTTTCTTTGAAAATTTTTTG
QY 711 ----- 710
  
```

```

Db 781 ATTGTAATCTCCGAACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 711 ----- 710
Db 841 ATATACGATATATGATGTGAAGAAACATGAATTTCCAGTATTTTAACCACTG
QY 711 --TGCTGACACGCTGACGAACAGCGCTGATTAACCTGAG----- 751
Db 901 ACGAACAACAAACATGACGAGAAAGAGATTAATCATGTGAAAGCTATTAAGAAC
QY 752 ----- 882
Db 961 GTGCTGCTACTCATCTGATCTGTTGCTGCAAGCTATTTAATATCATGACAGAAAG
QY 757 TAAATCATATTTGA-----CGTACCTTAAAGTATATCATGCGTAAATTTG
Db 1021 AAACAAACTGTGTGCTTCATTGATGTTGTTACCAACAGAAATTAAGAGTTAGTTG
QY 802 ACGATGAGATCTGTAATACACTCACTAATGAGGCGAATGGTAAACCGGCCCCCTG
Db 1081 AAGCATTAAGTCCCAAAATTTGTTTAAACATGTGATATCTTACTGATTTT
QY 862 AGGTGACGCTATCGATTAAC----- 882
Db 1141 CCAATGAGGCGACAGTTAAGCGCTAAAGGATTAATCCGCAATTTTAACTCT
QY 883 ----- 912
Db 1201 TCGAAGACAGAAATTTGCTGACATTTGTAATACATCAATTTGAGAGTCTG
QY 913 CCATAGTTTGAAGCGCGCC-----CACCAGGTGAGAGTCAAGCTT
Db 1261 TATACAGATATGCAATGAGGCGACATTAAGATGACACAGTGTGAGGCGCTAGTA
QY 957 TTTTCCCTTTAGTGAAGG-----TTAA 979
Db 1321 TTGTTAGCGGTTTGAAGGAGGCGGCAAGAAAGTAAACAAAGAACTTAAGAGCC
QY 980 TTGATCCATGCGGTGCAATTTTACGACATCTTCTTGAAGGTTAATCTAAGCTGATCA
Db 1381 TGTTAACAGAAATTTGTAAGGAGGCTCTTATCTAAGGAAATTAATCAAGGATCTG
QY 1040 GAATCATCTGCTGCG-----TCTTTTCCGCGCTCAATCAATCCGCAAGCTG
Db 1441 TTGACATTTGCAAGGCGCAAAAGATTTGTTATGCGCTTATGCTCAAAAGACATG
QY 1090 GCTATCTGAGCATCGGGAAGGAAAGCCCGTCTTTCCGCGAGTTGAAGCGCAT
Db 1501 GTGAAAGAGATGAAGTTTGAATTTGTTATTAATGAACCCGCTGAGGTTTAATGA
QY 1150 GGAAGAGATTTGCGAGAGATGACTGCTGCTGATTTGAAGGAGAAAGCAAGCTT
Db 1561 AGGAGAGCATGTTGGTCAACATATTAAGAACCTGATGATGATGCTTACAGAGCTG
QY 1210 CCATGATGATTCGGAAGGTTGCGCATATTTGCAAAAGGAGATGATGAAGGAT
Db 1621 ACATTTATTTTGTGGAAGAGCATATTTGCAAAAGGAGATGATGAAGGAT
QY 1255 AACTGTTGTTGAGGCACTGCGATTAACAGT-----CGTGGGCTTTCCGGAACA
Db 1681 AAGCTTACAGAAAGAGAGGCTGGGAAGCATTTTGAAGATGCGCGCAAACTTAA
QY 1310 GTTCCGATGCTGACCGCAAGCCATCAAGAACCGGAACATTAACGCGGAC-----
Db 1741 AAATGATATTAATTAATATCATGTATATCAAACTCAAAATTAAGCTTCAATTAAT
QY 1365 CGAAATCGCGTCCGCTGAGTGAATTAATGACAGGCTGCGCGCTGAGATTAAGCTC
Db 1801 TATATCAGTATTAATCCAAATTTCTATGTTGGCGCGCTGACAGAAATTCATTAAGCT
QY 1425 AGCGAGACGCGTATCTGCTGATGCGCAGAAATGAGATGATGATGATGATGATGAT
Db 1861 TATGATATCGTGAAGCTGAGGAGGCGCGCTGACCAAGCTTTTGTCTTATGATGAG
  
```


Db

4081 TTCCCCGAAAAGTGCAC 4098

RESULT 8
US-09-402-266B-18
Sequence 18, Application US/09402266B
Patent No. 6537767
GENERAL INFORMATION:
APPLICANT: HINNEN, Albert
APPLICANT: HEGEMANN, Johannes
APPLICANT: MCNDER, Thomas
APPLICANT: SCHUSTER, Tilmer
APPLICANT: KRAMER, Horst
APPLICANT: ZIMMERMANN, Friedrich
APPLICANT: ENTIAN, Karl-Dieter
TITLE OF INVENTION: Process for Screening Anticancerically Active Substances
FILE REFERENCE: 38005-0094
CURRENT APPLICATION NUMBER: US/09/402,266B
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: PCT/EP98/01904
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: DE 1973572.2
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.0
SEQ ID NO 18
LENGTH: 4088
TYPE: DNA
ORGANISM: Unknown
FEATURE:
NAME/KEY: misc feature
LOCATION: ()..71
OTHER INFORMATION: Synthetic sequence
US-09-402-266B-18

Query Match 62.9%; Score 2302; DB 4; Length 4088;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 550; Indels 426; Gaps 5;

QY 5 ATTGTAGCGCTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGTCAATT 64
Db 1 ATTGTAGCGCTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAGTCAATT 64
QY 65 TTTAACCAATAGCGCAATCGCAAAATCCCTTAATCAAAAGATAGACCGAGATA 124
Db 61 TTTAACCAATAGCGCAATCGCAAAATCCCTTAATCAAAAGATAGACCGAGATA 124
QY 125 GGGTTAGTGTGTCCTGATTGGAACAAGATCACTATTAAAGAGTGAATCCCAAC 184
Db 121 GGGTTAGTGTGTCCTGATTGGAACAAGATCACTATTAAAGAGTGAATCCCAAC 184
QY 185 GTCAAGAGGCGAAAAACCGTCTATCAGGGGATGCGCCACTACGTGAACATCACTCA 244
Db 181 GTCAAGAGGCGAAAAACCGTCTATCAGGGGATGCGCCACTACGTGAACATCACTCA 244
QY 245 TCAAGTTTTTGGGGTGAAGTCCGTTAAAGCACTAAATCGGAACCTTAAAGGAGCC 304
Db 241 TCAAGTTTTTGGGGTGAAGTCCGTTAAAGCACTAAATCGGAACCTTAAAGGAGCC 304
QY 305 CGATTTAGACTTGAAGCGGGAAGCGCGCAACGTGGCGAAGAAAGGAGGAGGAGCC 360
Db 301 CGATTTAGACTTGAAGCGGGAAGCGCGCAACGTGGCGAAGAAAGGAGGAGGAGCC 360
QY 365 AAAAGAGCGGCGCTAGAGCGCTGCAAGTATGAGCGTCAACGCTGCGGTAACCA 424
Db 361 AAAAGAGCGGCGCTAGAGCGCTGCAAGTATGAGCGTCAACGCTGCGGTAACCA 424
QY 425 CCCGCGCGCTTAATAGCGCGCTTCAAGAGCGCGTCCATTCGCGATTAGGCTGCGCA 484
Db 421 CCCGCGCGCTTAATAGCGCGCTTCAAGAGCGCGTCCATTCGCGATTAGGCTGCGCA 484

QY 485 TGTGGAGAGCGGATCGGTGCGGCTCTTGCCTATTACCCAGCTGGGGAAGGAGGA 544
Db 481 TGTGGAGAGCGGATCGGTGCGGCTCTTGCCTATTACCCAGCTGGGGAAGGAGGA 540
QY 545 TGTGTGCAAGCGGATTAAGTTGGGTAAACGCAAGGTTTCCAGTCAAGAGTGTAA 604
Db 541 TGTGTGCAAGCGGATTAAGTTGGGTAAACGCAAGGTTTCCAGTCAAGAGTGTAA 600
QY 605 ACAGAGCGGAGTGAAGCGGCTCTTGCCTATTACCCAGCTGGGGAAGGAGGA 664
Db 601 ACAGAGCGGAGTGAAGCGGCTCTTGCCTATTACCCAGCTGGGGAAGGAGGA 660
QY 665 ACTGCGGCGTAATGTTTACAGGATGAGG----- 700
Db 661 GGTGCGCTCTAGAACTAGTGAATCGGCGCAAGGAGGCTCTGTCAGAAATGACAGTATA 720
QY 701 --CAGATGAAGATGCTGACAGCTGCAAGACAGCACTAGATTAACTTAAAGATA 758
Db 721 GAATGATGATTAACCTTGTATCTTCAAGTATCACTAGTGTATATATTAATGACA 780
QY 759 ATGATATTGAGAGTACGTTAAAGATTAATCAATGCTAAATTAAGCATGAGATCTGTA 818
Db 781 TTCAATGATTAATCAATTAATCAATGATTAATTAATGCTGAGCTTTAAATTAATG 840
QY 819 TACGACTACTATAGGAGGAATGGGTACCGGCGGCGGCGGCGGCGGCGGCGG 878
Db 841 GTGTCACTAATTAAGAACACTTTGTGAGGAGGACATGCTGTGATCAATGGGAGGT 900
QY 879 AAGCTGATATGAAATTCGACAGCGCGGAGATCCAGTATGAG----- 927
Db 901 GAGCTTCTTTAGGCAACCGCAAGAGCTTGAACGACCTTACATCGATGATCAT 960
QY 928 -----CGGCGCGACCGGCGGAGCTCCAGCTTTGTTCCCTTATGAGGTTAAT 980
Db 961 CTTCCTTCGACAGCAATTAACAGTGAAGGATTTCTGATACCTTCGCAAGCTTTCA 1020
QY 981 TAGATCCATAGCTCAATTTTACGACACTATCTTTTCAAGGTTAATCTAGCTCA 1040
Db 1021 AAATGGGAGATCATTCGCAAGAGATCTTCTTCTTCTTCAACCAAGTTC 1080
QY 1041 GATCATATGTTGGGCTTTTTCGCTAGTATGATGCGCAAGCTGAGCTATCTGAG 1100
Db 1081 GACATGCTGTAACCGCTGTTGAAAGATCTACACGCTGGAAGGCTCTATCCA 1140
QY 1101 ATCGGAGAGGAAGCGGCGCTTTT-----CGGCGAGTTGAAGCGGATGGAAGGT 1158
Db 1141 AGCGCAATCTGATCCAAACCTTTTACTCCACGCGCAGTAGGCTCTTTAAAGC 1200
QY 1159 TTGCGAGAGTGAAGTCTGCTGATTGA----- 1186
Db 1201 TTGACGAGAGCAATCCGCACTTCAAGTGTGATGTCGTATGTTAGTATCACC 1260
QY 1187 ----- 1186
Db 1261 AATGACTCAACGATTAGGACCAACCGGAATCTTGGCCAGAGCATGATATATGTC 1320
QY 1187 CGTTGAGCGAAAAAGCACTTTACCATATGATTTGGAAGAGTGTGCGCATGACCTT 1246
Db 1321 CAGAAACCTTAATCTGTGTGAGATTAATCACTTGGATTGTGTGCTGCTAC 1380
QY 1247 TAAAGTGAATCTGTTGTTGAGCGCACTTGGGAATCAAGTGTGCGCTTTTCCGAG 1306
Db 1381 TGTCTTCTCTTTTCTGGAAGATCGAGTCTTATCGCTAAGGAGCAACCTTTAA 1440
QY 1307 ACAATTCGAGATGTTGAGCGCGAAGCAATCAAGCAATCCGGAATTAACGAGCA 1366
Db 1441 AGAGATGCAATCTGATCTTGTGTTTCAATTTGTAATGAGCTTACTAGGCTTTTCTC 1500
QY 1367 GAACTGCGGCGGCTGAGT----- 1391
Db 1501 TGTGATCTTGTCTTGTATCTTGTGCTGCTATTTTATGATTAATCTTGAAGAAAT 1560
QY 1392 ----- 1391

Dp	1561	CACATTACTTATATATAAATGATATATTCATATATGATATATCGGAATCGCTAAGAAAAA	1620
Qy	1392	-----	1391
Dp	1621	AAGAGTCATCGCTAGGGGAAAAAAAAAATGAAATCATTTACGAGGCAATAAAAATA	1680
Qy	1392	-----	1391
Dp	1661	TAGAGTATCTAGAGAGGCCAAGATATAGAAAAAATTCGGGAAAGACCTGTG	1740
Qy	1392	-----	1391
Dp	1741	TTATGACTTCCCTGACTATATGCCGTTCCTTCAACATTAACCTGCACTGCTCTTACGCTC	1800
Qy	1392	-----AATGACAGCGGTGCGCGCTGGGATATTAACGTACAGAGACGGGTA	1438
Dp	1801	ACCAAGCTCTTAAAAACGGCGCGCTGAGGAATTCGATATCAAGCTTATCGATACCGTCG	1860
Qy	1439	TCCTGGCTGAGATGCCGCAAAATGGAATGGAATACCCCGTAGATTACCCCGCGGCGCGC	1498
Dp	1861	ACCTTGAGAGGGGGGCGCGGTAACCCAGCTTTGTTCCTTTATGTAAGGTTAATTGCGGC	1920
Qy	1499	TTGGCGTATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACATTTCCA	1558
Dp	1921	TTGGGATATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACATTTCCA	1980
Qy	1559	CACAAATATCAGACCGGGAAGCATAAAGTGAAGCTGGGGGTGCTTAATGATGAGCTAA	1618
Dp	1961	CACAAATATCAGACCGGGAAGCATAAAGTGAAGCTGGGGGTGCTTAATGATGAGCTAA	2040
Qy	1619	CTCACATTAATTGGGTGCGCTCACTGCGGCTTCCAGTGGGGAACCTGTCCGTCCAG	1678
Dp	2041	CTCACATTAATTGGGTGCGCTCACTGCGGCTTCCAGTGGGGAACCTGTGTGTGCGAG	2100
Qy	1679	CTGCATTAATGATCGGCCAAACGCGCGGGGAGAGCGGTTTGCATTTGGCGCTCTTCC	1738
Dp	2101	CTGCATTAATGATGATGGGCCAAACGCGCGGGGAGAGCGGTTTGCATTTGGCGCTCTTCC	2160
Qy	1739	GCTTCTCTGCTCATCTGACTCGCTGCGCTCGGTCTGTGGCTGCGGCGAGCGGATACACT	1798
Dp	2161	GCTTCTCTGCTCATCTGACTCGCTGCGCTCGGTCTGTGGCTGCGGCGAGCGGATACACT	2220
Qy	1799	CACTCAAAGGGGGTAAATACGGTTATCCACAAATCAGGGGATTAACGACGGAATAAATG	1858
Dp	2221	CACTCAAAGGGGGTAAATACGGTTATCCACAAATCAGGGGATTAACGACGGAATAAATG	2280
Qy	1859	TGAGCAAAAGGCCAGCAAAAGGCCAGAAACGCTAAAAAGGCCGGGTGCTGGCTTTTTC	1918
Dp	2281	TGAGCAAAAGGCCAGCAAAAGGCCAGAAACGCTAAAAAGGCCGGGTGCTGGCTTTTTC	2340
Qy	1919	CATAGAGCTCGGCCCCCTGACGAGCATCAAAAATCGACGCTCAGAGTGCGA	1978
Dp	2341	CATAGAGCTCGGCCCCCTGACGAGCATCAAAAATCGACGCTCAGAGTGCGA	2400
Qy	1979	AACCGGACAGAGCTATTAAGATACAGGCGTTTCCCTTGGAACTCCCTGTGCGGCTC	2038
Dp	2401	AACCGGACAGAGCTATTAAGATACAGGCGTTTCCCTTGGAACTCCCTGTGCGGCTC	2460
Qy	2039	CCTGTTCGACCCCTGCGCTTACCGGAATCCTGTCCGCTTTCCTTCCTTGGGAAAGGTG	2098
Dp	2461	CCTGTTCGACCCCTGCGCTTACCGGGAATGCGGCTTTCCTTCCTTGGGAAAGGTG	2520
Qy	2099	GCGCTTTCATAGCTCACGCTGTAGATATCTCAGTTGGGTAGAGTGCTTCCGTCAG	2158
Dp	2521	GCGCTTTCATAGCTCACGCTGTAGATATCTCAGTTGGGTAGAGTGCTTCCGTCAG	2580
Qy	2159	CTGGAGCTGTGACAGAAACCCCGGTTACGCGCGGACGCGCTGAGGCTTATCCGGTAATAT	2218
Dp	2581	CTGGAGCTGTGTGACAGAAACCCCGGTTACGCGCGGACGCGCTTATCCGGTAATAT	2640
Qy	2219	CGTCTTAGTCCAAACCGGTAAGACGACTTATCGCCATCGGACGACCACTGTGTAC	2278

Db	2641	CGTCTTGAAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGCGACAGCAACTGCTAAAC	2700
Qy	2279	AGGATTACAGAGCGAGGTATGTAGCGCGTGTCTACAGAGTTCTTTGAATGCTGGCCCTAAC	2338
Db	2701	AGGATTACAGAGCGAGGTATGTAGCGCGTGTCTACAGACTTCTTTGAATGCTGGCCCTAAC	2760
Qy	2339	TACGGCTTCACTAAGAAAGACAGTATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC	2398
Db	2761	TACGGCTTCACTAAGAAAGACAGTATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC	2820
Qy	2399	GGAAAAAAGATTGGAGCTCTTGATCCGGGCAAAACAAACACCGGTGATGGCGGTGTTTT	2458
Db	2821	GGAAAAAAGATTGGAGCTCTTGATCCGGGCAAAABACACCGGTGATGGCGGTGTTTT	2880
Qy	2459	TTTGTGTTGCAAGCAGACAGATTACCGCGAGAAAAAAGATCTCAAGAAAGATCCTTTGATC	2518
Db	2881	TTTGTGTTGCAAGCAGACAGATTACCGCGAGAAAAAABAGATCTCAAGAAAGATCCTTTGATC	2940
Qy	2519	TTTTCTACGGGGTCTGACGCTCACTGGAACGAAACTCAAGTTAAGGAAATTTTGGTCATG	2578
Db	2941	TTTTCTACGGGGTCTGACGCTCACTGGAACGAAAACTCAAGTTAAGGAAATTTTGGTCATG	3000
Qy	2579	AGATTATCAAAAAGATCTTCAAGCTATACCTTTAATTTAATAATGAAGTTTAAATCA	2638
Db	3001	AGATTATCAAAAAGATCTTCAAGCTATACCTTTAATTTAATAATGAAGTTTAAATCA	3060
Qy	2639	ATCTTAAGATATATATGAGTAAACTTGGCTGACAGATTACCAATGCTTATACAGTAGACA	2698
Db	3061	ATCTTAAGATATATATGAGTAAACTTGGCTGACAGTTATCCAATGCTTATACAGTAGACA	3120
Qy	2699	CCATATCTCAGGAGATCTGTCTATTTTCGTTCACTCAATAGTTGCTGACCTCCCGCTGTGTAG	2758
Db	3121	CCATATCTCAGGAGATCTGTCTATTTTCGTTCACTCAATAGTTGCTGACCTCCCGCTGTGTAG	3180
Qy	2759	ATTAACTAAGATACGGGAGGGCTTACATCTGCGCCCACTGTCTGAATGATATCCGCGAGAC	2818
Db	3181	ATTAACTAAGATACGGGAGGGCTTACATCTGCGCCCACTGTCTGAATGATATCCGCGAGAC	3240
Qy	2819	CCAGGCTCACCGGGCTCCAGATTTTATTCAGCAATAAACAGCAGCGGAGGCGCGAGACGC	2878
Db	3241	CCAGGCTCACCGGGCTCCAGATTTTATTCAGCAATAAACAGCAGCGGAGGCGCGAGACGC	3300
Qy	2879	AGAAATGGTCCCTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCT	2938
Db	3301	AGAAATGGTCCCTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCT	3360
Qy	2939	AGAGTAAGTATGCGCAATTAATATGTTTGGCAACGTTGTCGATGCTTACAGGCATC	2998
Db	3361	AGAGTAAGTATGCGCAATTAATATGTTTGGCAACGTTGTCGATGCTTACAGGCATC	3420
Qy	2999	GTTGGTGTACAGCTGTGTGGTTTGGATATGGCTTCAATACGCTCCGGTTCGCAACCATCAAG	3058
Db	3421	GTTGGTGTACAGCTGTGTGGTTTGGATATGGCTTCAATACGCTCCGGTTCGCAACCATCAAG	3480
Qy	3059	CGAGTTACATGATCCCCATGTTGTGTGCAAAAAAGCGTTAGTCTCTCGCTCCGCATC	3118
Db	3481	CGAGTTACATGATCCCCATGTTGTGTGCAAAAAAGCGTTAGTCTCTCGCTCCGCATC	3540
Qy	3119	GTTGTCAAGAAATGTTGGCGCGAGTTTATCACTCATAGTGTATGGCAGCATGCAATAT	3178
Db	3541	GTTGTCAAGAAATGTTGGCGCGAGTTTATCACTCATAGTGTATGGCAGCATGCAATAT	3600
Qy	3179	TCTCTTATCTGTCAAGCATCCGTAAGATGCTTTTCTGTGATCTGTGATGATCTCAACCAAG	3238
Db	3601	TCTCTTATCTGTCAAGCATCCGTAAGATGCTTTTCTGTGATCTGTGATGATCTCAACCAAG	3660
Qy	3239	TCATTTCAAGAAATGTTATGCGCGCGACGAGTTGCTTTGCCCGGCTCAATATACGGAAT	3298
Db	3661	TCATTTCAAGAAATGTTATGCGCGCGACGAGTTGCTTTGCCCGGCTCAATATACGGAAT	3720
Qy	3299	AATACCGCGCCATACAGAACTTTAAAGTGTCTATCATTTGGAAAAAGCTTTCTTGGGG	3358
Db	3721	AATACCGCGCCATACAGAACTTTAAAGTGTCTATCATTTGGAAAAAGCTTTCTTGGGG	3780

```

QY      3359  CGAAAACCTCCAGAGATCTTACCGCTGTGAGATCCAGTGTAAATCCCACTCGTGCA 3418
      |||
Db      3781  CGAABACTCTCAGAGATCTTACCGCTGTGAGATCCAGTGTAAATCCCACTCGTGCA 3840
QY      3419  CCCCACTGATCTTCAGCATCTTTTACTTTCACACAGCGTTCTGTGGTAGCAAAAACAGGA 3478
      |||
Db      3841  CCCCACTGATCTTCAGCATCTTTTACTTTCACACAGCGTTCTGTGGTAGCAAAAACAGGA 3900
QY      3479  AGGCAAAATGCGCGCAAAAAGGAAATTAAGGCGCACAGAAATGTTGAATCTCATCTC 3538
      |||
Db      3901  AGGCAAAATGCGCGCAAAAAGGAAATTAAGGCGCACAGAAATGTTGAATCTCATCTC 3960
QY      3539  TTCTTTTCAATATTATTGAAGCATTTATCAGGTTATTTGTCTATGAGCGGATACATA 3598
      |||
Db      3961  TTCTTTTCAATATTATTGAAGCATTTATCAGGTTATTTGTCTATGAGCGGATACATA 4020
QY      3599  TTGATGTATTGAAAAAATTAACCAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTG 3658
      |||
Db      4021  TTGATGTATTGAAAAAATTAACCAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTG 4080
QY      3659  CCAC 3662
      |||
Db      4081  CCAC 4084

```

RESULT 9

```

US-09-993-170-1
; Sequence 1, Application US/09993170
; Patent No. 6743969
; GENERAL INFORMATION:
; APPLICANT: Valent, Barbara
; APPLICANT: Bryan, Gregory
; TITLE OF INVENTION: Modification of a Plant Disease Resistance Gene
; TITLE OF INVENTION: Specificity and Method for Engineering Altered
; FILE REFERENCE: BBI385 US NA
; CURRENT APPLICATION NUMBER: US/09/993,170
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/248,335
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-09-993-170-1

```

```

Query Match      62.2%; Score 2279.2; DB 4; Length 4119;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 3151; Conservative 0; Mismatches 508; Indels 457; Gaps 11;

```

```

QY      1  CTAATTGTAAAGCTTAATATTTTGTAAATTCGGCTTAAATTTTGTAAATCAGCTC 60
      |||
Db      4  CTAATATGTAAGCTTAATATTTTGTAAATTCGGCTTAAATTTTGTAAATCAGCTC 63
QY      61  ATTTTAAACCAATAGGCGGAAATGCGCAAAATCCCTATTAATCAAAAAGATAGACGA 120
      |||
Db      64  ATTTTAAACCAATAGGCGGAAATGCGCAAAATCCCTATTAATCAAAAAGATAGACGA 123
QY      121  GATAGGTTGAGTGTTCAGATTGGAACAAGATCCACTATTAAGAAGCTGAGATC 180
      |||
Db      124  GATAGGTTGAGTGTTCAGATTGGAACAAGATCCACTATTAAGAAGCTGAGATC 183
QY      181  CAAAGTCAAAGGCGCAAAAACGCTCTATCAGGCGATGGCCCACTAGGAACATCAAC 240
      |||
Db      184  CAAAGTCAAAGGCGCAAAAACGCTCTATCAGGCGATGGCCCACTAGGAACATCAAC 243
QY      241  CTAATCAGTTTTTTGGGCTGAGGTGCGCTAAAGACTAAATCGAAACCTTAAAGGAG 300
      |||

```

```

Db      244  CTAATCAGTTTTTTGGGCTGAGGTGCGCTTAAAGCACTAAATCGGAACCTTAAAGGAG 303
QY      301  CCCCGATTTAAGCTTTAAGCGGGAAAGCCGCGCAACGTGGCGAAGAAAGGAGAGA 360
      |||
Db      304  CCCCGATTTAAGCTTTAAGCGGGAAAGCCGCGCAACGTGGCGAAGAAAGGAGAGA 363
QY      361  AGGAAAGAGAGCGGCGCTAGAGGCGTGAAGTGTAGCGTCAAGCTGCGTAAACAC 420
      |||
Db      364  AGGAAAGAGAGCGGCGCTAGAGGCGTGAAGTGTAGCGTCAAGCTGCGTAAACAC 423
QY      421  CACACCGCGCGGCTTAATGCGCGCTTACAGGCGCGCTCCATTTGCGCATCAGCTGCG 480
      |||
Db      424  CACACCGCGCGGCTTAATGCGCGCTTACAGGCGCGCTCCATTTGCGCATCAGCTGCG 483
QY      481  CAATCTTTGGAAGGCGATTCGGTGGGATACGCGAGGTTTCCAGTCAAGCTGCGAAAG 540
      |||
Db      484  CAATCTTTGGAAGGCGATTCGGTGGGATACGCGAGGTTTCCAGTCAAGCTGCGAAAG 543
QY      541  GGGATGTGTCGAAGGCGATTAAGTTGGGTACGCGAGGTTTCCAGTCAAGCTGCG 600
      |||
Db      544  GGGATGTGTCGAAGGCGATTAAGTTGGGTACGCGAGGTTTCCAGTCAAGCTGCG 603
QY      601  TAAAGCAAGCGCACTGA-----GCGGCGCTGCTT 630
      |||
Db      604  TAAAGCAAGCGCACTGAATTTGTAATGACATCACTATAGGCGCAATTTGGTTACGCGGCC 663
QY      631  CATTACGTTTTTGAACCCGTGAGAGAGCGGCACTCGCGGTGCAAAATGTTTTCACG 690
      |||
Db      664  CCCCCTCGAGGTGCATCGAGGTAGGTCTAGGGGCTGATCTCACAATATTTTGTGA 723
QY      691  CGTATGAGAGATGAAGTCTCGACAGCGTGCAGAACCGCACT-----728
      |||
Db      724  ATTTCAAAATGCGAGGAGCATGAATATGATGATTTAATATATGGCCCACTTAC 783
      |||
QY      739  -----AATTAACCTTGAAGAAATATCATTTGTGAGCTACGTTAAGATTAATCAT 790
      |||
Db      784  CTTATATTAATTTGTGATGATATCTAATTAATGAGCTAATTTACGCTAATTAAT 843
      |||
QY      791  GCGTAAATTTGACGATGGGATCTGTATACAGCTCACTA--TAGGCGAATGGGTAC 848
      |||
Db      844  GCAGATTAATTAACATACGAAATTAATTCGATTAATTAATTAATTAATTAATTAATTA 903
      |||
QY      849  GGGCCCCCCTCGAGGTGACGATGATCAATGATTAATTAATTAATTAATTAATTAATTA 908
      |||
Db      904  TAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 963
      |||
QY      909  GGATCAGTATTTAGAGGCGCGCACGCGGTGAGCTCAGCTTTGTCCCTTTA 968
      |||
Db      964  TAAATTAAGTTGGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1023
      |||
QY      969  GTGAG-----GTTAATTAGATCCATGCTCAATTTTACGAGACTATCTTCT 1018
      |||
Db      1024  TAGAGAACTTTGCTTAATTAATTTCCAAACATTAACAAATTAATCACTGACTGAAATG 1083
      |||
QY      1019  AGGTTATATCTTACCTCAGATCATTA-----1047
      |||
Db      1084  ACAATTTGTTTGAACAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1047
      |||
QY      1048  -----TCGCGGCTTTTTCGCGCTCAGTATGCGCAAGCTGGCGGTATTTGGG 1099
      |||
Db      1144  TTAATTTTGCATCGATGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1203
      |||
QY      1100  CAT-----1102
      |||
Db      1204  CTTTACCAACATTTGGAACCTTTTACACCCAGTTTACATTAATCAATCAATTCCAAC 1263
      |||
QY      1103  -----CGGGAAGAGAGAGCCGCGCTTT-----1127
      |||
Db      1264  ATATCAAGGAGATTTAAGAGCGGCTTTATTTGAACGCTATTTCCAAATTTCAAGATT 1323
      |||
QY      1128  -----TCCCGAGGTTTGAAGCGGATGGAAGAGTGTTCGCGAGATGACTGCTGTGCA 1182
      |||
Db      1324  CGCAGGCTTCGAAATTCGTGCGCGCTTAAGAAAGTTGTGCGAGCTGCTGTGGGGCT 1383
      |||

```

1183 TTGAGCTTGAGCGAAACGCA--CGTTTACATGATGATGGGAG----- 1227
1384 ATCAAGCGCGTTAAATGACATCGTTATTTAGATTAATCTTTAAACGACAGCAG 1443
1228 ----- 1227
1444 ATATTCAAACTGGGCTTCAAAAAGATTTTAAAGAAATTTACAGAAATGTAAAGAGGACG 1503
1228 -----GTGTGCCATGACAGCGCTTTAACGGTAACTGTGTTCAGGCACT 1275
1504 CGGAGAAATTTCTTAACCTGCGCAGATTAATAATTTATACGTGCGCTCGAGAAAGAG 1563
1276 GGGATCACAGTTGCTGCGGCTTTTCCGACACAGTTCCGATGTCAAGCCCGAAGCGCA 1335
1564 TTCAATATTTGGCGATGACCTTATTTAAAGAAAAGAAATTTATATATGCTCTCTTCT 1623
1336 TCAGCAACCCGGAACAAATACCGGCGACA----- 1362
1624 TCACAAACCCGTTAAACAGCAGGAAATTAATGCGGTTAACCAAGATACATTAATTTAC 1683
1363 -----GCCGGAATGCCGTGCCGTG 1383
1584 ATGAAATGTGCATATAATTTTAAAGAGTGAAAGATTTATGTATACAAATGGGATGGGA 1743
1384 TGCAGATTAATGACAGCGGTCGGCGCTGGGATATTAGTACGCGAGAGCGGATTCCTG 1443
1744 TTCAAAATTTGATAGTACAAAGATTTAAACCCCGACAGTTATCTATTTTGTGAC 1803
1444 GCTGGATGCCGAGAAATGACATGATATCCCGTGAATTAACCGGCGGG----- 1493
1804 AATGTGACGTTATTAATATTGTTAAATACGTTATGTTGAATGAGGAAATCGCGGT 1863
1494 ----- 1493
1864 CAATCAACATAGAGGATCACTAGTTCTAGAGCGGCGCACCGCGGTGAGACTCA 1923
1494 -----CGCGTTGGCGTAACTAGTGCATAGCTGT 1523
1924 GCTTTGTTCCCTTATGAGGAGTTAATTTGAGGTTGGCGTAAATCAATGATCATAGCTGT 1983
1524 TTCTGTGTGAATTTGTTATCCGCTCACAAATTCACACAACTACAGACCGGAAAGCATTA 1583
1984 TTCTGTGTGAATTTGTTATCCGCTCACAAATTCACACAACTACAGACCGGAAAGCATTA 2043
1584 AGTGTAAAGCTTGGGGTGCCTTAATGATGAGCTAACTCAATTAATTCGTTGGCTCAC 1643
2044 AGTGTAAAGCTTGGGGTGCCTTAATGATGAGCTAACTCAATTAATTCGTTGGCTCAC 2103
1644 TGCCCGCTTTTCACTCGGGAACCTGTGTGTCAGCTGCATTAATGAATGGCCAAACGCG 1703
2104 TGCCCGCTTTTCACTCGGGAACCTGTGTGTCAGCTGCATTAATGAATGGCCAAACGCG 2163
1704 CGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCGCGCTTCTCGCTCACTGACTCGCTGC 1763
2164 CGGGGAGAGCGGTTTGGCTATTGGGCGCTCTTCGCGCTTCTCGCTCACTGACTCGCTGC 2223
1764 GCTCGGTGTTGCGCTGCGGAGCGGATACGCTCACTCAAAAGCGGTTAAACGTTAT 1823
2224 GCTCGGTGTTGCGCTGCGGAGCGGATACGCTCACTCAAAAGCGGTTAAACGTTAT 2283
1824 CCACAGAAATCAGGGGATTAACCCAGAAAGAAATGTGACAAAGGCGCAAGAAAGCGCA 1883
2284 CCACAGAAATCAGGGGATTAACCCAGAAAGAAATGTGACAAAGGCGCAAGAAAGCGCA 2343
1884 GGAACCGTTAAAGGCGGTTGCTGGGCTTTTTCATAGGCTCGGCGCCCTGACGAGC 1943
2344 GGAACCGTTAAAGGCGGCTTGTGCTGCTTTTTCATAGGCTCGGCGCCCTGACGAGC 2403
1944 ATCAAAATTCAGCTCAAGTACAGAGTGGCGAAACCCGACAGAGCTATTAAGATACC 2003
2404 ATCAAAATTCAGCTCAAGTACAGAGTGGCGAAACCCGACAGAGCTATTAAGATACC 2463

2004 AGGCGTTTCCCGCTGGAAGCTCCCTGAGGCTCTCCGTTCCGACCCCTGCGGCTTACG 2063
2464 AGGCGTTTCCCGCTGGAAGCTCCCTGAGGCTCTCCGTTCCGACCCCTGCGGCTTACG 2523
2064 GATACCTGTCCGCTTCTCCCTTCGGGAACGTGCGCTTCTCATAGTACAGCTGTA 2123
2524 GATACCTGTCCGCTTCTCCCTTCGGGAACGTGCGCTTCTCATAGTACAGCTGTA 2583
2124 GGTATCTCAGTTGCGGTAGTGGTCTGCTCAAGCTGGGCTGTGTGACGAACCCCGG 2183
2584 GGTATCTCAGTTGCGGTAGTGGTCTGCTCAAGCTGGGCTGTGTGACGAACCCCGG 2643
2184 TTCAGCCGACCGGCTGAGCTTATCCGTTAATTCGTTGAGTCCAAACCCGTTAAAC 2243
2644 TTCAGCCGACCGGCTGAGCTTATCCGTTAATTCGTTGAGTCCAAACCCGTTAAAC 2703
2244 ACGATTAATCGCACCTGGCAGACCACTGGTAAACAGATTAAGACAGAGATATGAG 2303
2704 ACGATTAATCGCACCTGGCAGACCACTGGTAAACAGATTAAGACAGAGATATGAG 2763
2304 GCGGTGCTACAGAGTTCTTGAAAGTGTGCTTAACTACGCTTACAGTAAGAGACATAT 2363
2764 GCGGTGCTACAGAGTTCTTGAAAGTGTGCTTAACTACGCTTACAGTAAGAGACATAT 2823
2824 TTGATATCTGCGCTCTGCTGAGAGCAGTTAATCTTGGAAAAAGAGTTGATCTTGAT 2883
2424 CCGGCAAAACAAACACGCTGTGAGCGGTGTTTGTGTTGTTGAAAGCAGAGATTACG 2483
2884 CCGGCAAAACAAACACGCTGTGAGCGGTGTTTGTGTTGTTGAAAGCAGAGATTACG 2943
2484 GCAGAAAAAGATCTCAAGAGATCTTGTATCTTTTCTACGCGGCTGACGCTCAGT 2543
2944 GCAGAAAAAGATCTCAAGAGATCTTGTATCTTTTCTACGCGGCTGACGCTCAGT 3003
2544 GGAACGAAATCTACGTTAAGGGAATTTGTGATGAGATTAATCAAAAAGATCTTCACT 2603
3004 GGAACGAAATCTACGTTAAGGGAATTTGTGATGAGATTAATCAAAAAGATCTTCACT 3063
2604 AGATCTTTTAAATTAATAATGAATTTAAATCAATCTAAAGATATATAGTAAACTT 2663
3064 AGATCTTTTAAATTAATAATGAATTTAAATCAATCTAAAGATATATAGTAAACTT 3123
2664 GGTCTGACAGTTACCAATGCTTAATCAGTAGGACCTATCTCAGCATCTGTCTATTC 2723
3124 GGTCTGACAGTTACCAATGCTTAATCAGTAGGACCTATCTCAGCATCTGTCTATTC 3183
2724 GTTATCTCAATGTTGCTGACTCCCGTCCGTGTGATGATTAATCAATACGGAAGGCTTAC 2783
3184 GTTATCTCAATGTTGCTGACTCCCGTCCGTGTGATGATTAATCAATACGGAAGGCTTAC 3243
2784 CATCTGCCCCAGTGTGCAATGATTAACGAGAACCCAGCTCAACGCGCTCAAGATTAT 2843
3244 CATCTGCCCCAGTGTGCAATGATTAACGAGAACCCAGCTCAACGCGCTCAAGATTAT 3303
2844 CAGCAATTAACCAAGCCGGAAGGCGCAGACGCAAGAGTGTCTGCAACTTTATCCG 2903
3304 CAGCAATTAACCAAGCCGGAAGGCGCAGACGCAAGAGTGTCTGCAACTTTATCCG 3363
2904 CCTTCATCCAGTCTAATTAATTTGTTGCGGGAAGCTAAGTAAGTATGCTGCACTTAATA 2963
3364 CCTTCATCCAGTCTAATTAATTTGTTGCGGGAAGCTAAGTAAGTATGCTGCACTTAATA 3423
2964 GTTGGCAACAGTTGTTGCACTTGTCAAGGATCGGTGTGCAAGCTGCTGTTGGTA 3023
3424 GTTGGCAACAGTTGTTGCACTTGTCTACAGGATCGGTGTGCAAGCTGCTGTTGGTA 3483
3024 TGCTTCAATTCAGCTCGGTTCCCAAGATCAAGGAGATTAATGATCCCAATGTTGT 3083
3484 TGCTTCAATTCAGCTCGGTTCCCAAGATCAAGGAGATTAATGATCCCAATGTTGT 3543
3084 GCAGAAAAAGGCTTACGCTTCCGCTCTCGATCTGTTGTCAAGATTAAGTGGCCGAC 3143

Accession	Sequence	Length
Db	3544 GCAAAAAGGGGTAGCTTCCTTCGCTCCGATCGTTGCACAAATTAAGTGGCCGAG	3603
Qy	3144 TGTATCACTCATGGTTATAGCAGCACTGCATATATCTCTACTGTACATCCGTAA	3203
Db	3604 TGTATCACTCATGGTTATAGCAGCACTGCATATATCTCTACTGTACATCCGTAA	3663
Qy	3204 GATGCTTTTCGTGACGTGTAGTACTCCAAAGTATCTGAGATAGTGTATGCGGC	3263
Db	3664 GATGCTTTTCGTGACGTGTAGTACTCCAAAGTATCTGAGATAGTGTATGCGGC	3723
Qy	3264 GACCGATGTCCTTTCGCCGCGCTCATACGGGATTAATACGGCGCATATGCAAACTT	3323
Db	3724 GACCGATGTCCTTTCGCCGCGCTCATACGGGATTAATACGGCGCATATGCAAACTT	3783
Qy	3324 TAAAGTCTCATATTTGGAACAGTTCTTCGCGCGGAAACCTTCAGAGATCTTAACGC	3383
Db	3784 TAAAGTCTCATATTTGGAACAGTTCTTCGCGCGGAAACCTTCAGAGATCTTAACGC	3843
Qy	3384 TGTGAGATCCAGTTCGATGTAACCACTCGTGACCCACATGATCTTCAGCATCTTTTA	3443
Db	3844 TGTGAGATCCAGTTCGATGTAACCACTCGTGACCCACATGATCTTCAGCATCTTTTA	3903
Qy	3444 CTTTCACACAGGTTTCGCGGTGAGCAAAAACAGGAAGCAAAATGCCCAAAAAAGGAA	3503
Db	3904 CTTTCACACAGGTTTCGCGGTGAGCAAAAACAGGAAGCAAAATGCCCAAAAAAGGAA	3963
Qy	3504 TAAAGCGACACGGAATGTTGAATACTCATCTTCTCTTTCATATATTGAAAGA	3563
Db	3964 TAAAGCGACACGGAATGTTGAATACTCATCTTCTCTTTCATATATTGAAAGA	4023
Qy	3564 TTTATCAGGGTTATTTGTCTCATAGACCGGATCATATTTGAATGATTTGAAAAATTAAC	3623
Db	4024 TTTATCAGGGTTATTTGTCTCATAGACCGGATCATATTTGAATGATTTGAAAAATTAAC	3683
Qy	3624 AATTAAGGGTTCCGCGCACATTTCCCGGAAAAAGTGC	3659
Db	4084 AATTAAGGGTTCCGCGCACATTTCCCGGAAAAAGTGC	4119

Query Match	59.3%	Score 2172.6;	DB 3;	Length 7102;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 2177; Conservative	0;	Mismatches	7;	Indels 0;

1479 GAGTTACCCGCGCGCGCTTGCGTATCATGTGATAGCTGTTTCTGTGGAATT 1538

Dp	3341	GAGTCGACCGGCATGCAAGCTTGGCGATATCATGCTGTCATAGCTGTTTCTGTGTGAATTT	3400
Oy	1539	GTTATCCGCTCACAATTCACAACAATACGAGCCGGAGACATAAAGTGTAAAGCTCGG	1598
Dp	3401	GTTATCCGCTCACAATTTCCACAACAATACGAGCCGGAGACATAAAGTGTAAAGCTCGG	3460
Oy	1599	GTCCTAATAGTAGAGCTAACTCACTAATTAATGCGTTGGGCTCACTGCCCCCTTTCAGT	1658
Dp	3461	GTCCTAATAGTAGAGCTAACTCACTAATTAATGCGTTGGGCTCACTGCCCCCTTTCAGT	3520
Oy	1659	CGGAAAACCTGTGTGTCCAGCTGACATTAAATGAAATCGGCCACGCGGGGAGAGCGGTT	1718
Dp	3521	CGGAAAACCTGTGTGTCCAGCTGACATTAAATGAAATCGGCCACGCGGGGAGAGCGGTT	3580
Oy	1719	TGCGTATTTGGGCGCTTTCCGCTCTCTCGCTCACTGACTCGTGGCTCGTTCGTGGC	1778
Dp	3581	TGCGTATTTGGGCGCTTTCCGCTCTCTCGCTCACTGACTCGTGGCTCGTTCGTGGC	3640
Oy	1779	TGCGGAGACGGTATACAGCTCACTGAAAAGCGGTAAATCGGTTATCCAGAAATCAGGG	1838
Dp	3641	TGCGGAGACGGTATACAGCTCACTGAAAAGCGGTAAATCGGTTATCCAGAAATCAGGG	3700
Oy	1839	ATTAAGCAGAAAAGAAATGTGAGCAAAAGCCACGAAAGGCCAGAAACCGTAAAAAG	1898
Dp	3701	ATTAAGCAGAAAAGAAATGTGAGCAAAAGCCACGAAAGGCCAGAAACCGTAAAAAG	3760
Oy	1899	CCGCGTTGCTGCGCTTTTTCATAGGCTCGCCCCCTGACGAGCATCAAAAATTCGAC	1958
Dp	3761	CCGCGTTGCTGCGCTTTTTCATAGGCTCGCCCCCTGACGAGCATCAAAAATTCGAC	3820
Oy	1959	GCTCAAGCAGAGGTGGGAAACCCGACAGAGCTATTAAGATACGAGCGTTTCCCTCTG	2018
Dp	3821	GCTCAAGCAGAGGTGGGAAACCCGACAGAGCTATTAAGATACGAGCGTTTCCCTCTG	3880
Oy	2019	GAACTCCCTGTGTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACGTTCGGCT	2078
Dp	3881	GAACTCCCTGTGTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACGTTCGGCT	3940
Oy	2079	TTCTCCCTTTCGGGAAAGGTGGCGCTTTCTCATAGCTCAAGCTGTAGTATCTCAGTTCCG	2138
Dp	3941	TTCTCCCTTTCGGGAAAGGTGGCGCTTTCTCATAGCTCAAGCTGTAGTATCTCAGTTCCG	4000
Oy	2139	TGTAGTCTGTTGCTCCAAAGCTGGGCTGTGTGACGAACCCCGTTCAGCCCGACCGT	2198
Dp	4001	TGTAGTCTGTTGCTCCAAAGCTGGGCTGTGTGACGAACCCCGTTCAGCCCGACCGT	4060
Oy	2199	GGCCTTATCCGGTAACTATGCTTTGAGTCCAACCCGGTAAAGACAGACTTATGGCAC	2258
Dp	4061	GGCCTTATCCGGTAACTATGCTTTGAGTCCAACCCGGTAAAGACAGACTTATGGCAC	4120
Oy	2259	TGGCAGACGCACTGTGTAAACAGATTTACAGAGCGAGGTATGTAGCGGCTGTACAGAT	2318
Dp	4121	TGGCAGACGCACTGTGTAAACAGATTTACAGAGCGAGGTATGTAGCGGCTGTACAGAT	4180
Oy	2319	TCTTGAAGTGGGCGCTAACTTACGGCTACCTTAAGAGCAGTATTTGTATCTGCGCTC	2378
Dp	4181	TCTTGAAGTGGGCGCTAACTTACGGCTACCTTAAGAGCAGTATTTGTATCTGCGCTC	4240
Oy	2379	TGCTGAAGCAGTATCTTCGAAAAAAGATTTGTAGTCTTGAATCCGGCAACAAACCA	2438
Dp	4241	TGCTGAAGCAGTATCTTCGAAAAAAGATTTGTAGTCTTGAATCCGGCAACAAACCA	4300
Oy	2439	CGGCTGGTGAAGGCGGTATTTTGTTCGACAGCAGATTTACGCGCAGAAAAAAGAT	2498
Dp	4301	CGGCTGGTGAAGGCGGTATTTTGTTCGACAGCAGATTTACGCGCAGAAAAAAGAT	4360
Oy	2499	CTCAAGAAAGATCCCTTGTATCTTTCTTACAGGAGCTGACGCTCAGTGAACGAAACTCAC	2558
Dp	4361	CTCAAGAAAGATCCCTTGTATCTTTCTTACAGGAGCTGACGCTCAGTGAACGAAACTCAC	4420
Oy	2559	GTTAAGGATTTTGTGCATGAGATTAACAAAAGATCTTCACTGATCTTTTAAATT	2618
Dp	4421	GTTAAGGATTTTGTGCATGAGATTAACAAAAGATCTTCACTGATCTTTTAAATT	4480

[illegible]

OY	2559	GTAAAGGATTTTGGTCATGATTA	CAAAAAGGATCTTCACAGATCCTTTAAAT	2618
DB	4421	TTTTTTT		

321 G11AAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATT 4480

QY 2079 TTCTCCCTTGGGAAAGCGTGGCGCTTTCTAGAGCTCAGCTGATAGTATCTCACTTGG 2138
 DB 3941 TTCTCCCTTGGGAAAGCGTGGCGCTTTCTAGAGCTCAGCTGATAGTATCTCACTTGG 2138
 QY 2139 TGTAGGTGTTGCTGCAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCT 4000
 DB 4001 TGTAGGTGTTGCTGCAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCT 4000
 QY 2199 GCGCTTATCCGGGAACTATGCTGCTTGAATCCAAACCCGGTAAGACAGCACTTATGCCAC 4060
 DB 4061 GCGCTTATCCGGGAACTATGCTGCTTGAATCCAAACCCGGTAAGACAGCACTTATGCCAC 4060
 QY 2259 TGGCAGCAGCCACTGTGTAAACAGGATTAGCAGAGCGAGTATGTAGCGGTCTACAGAGT 4120
 DB 4121 TGGCAGCAGCCACTGTGTAAACAGGATTAGCAGAGCGAGTATGTAGCGGTCTACAGAGT 4120
 QY 2319 TCTTGAAGTGTGGGCTTAACTACGCGTACACTAGAGAGCAGTATTTGTATCTGGGCTC 4180
 DB 4181 TCTTGAAGTGTGGGCTTAACTACGCGTACACTAGAGAGCAGTATTTGTATCTGGGCTC 4180
 QY 2379 TGTGAGGCAAGTTACCTTCCGAAAAAGTTGGTGTGATCTTGAATCCGCAAAACAAACA 4240
 DB 4241 TGTGAGGCAAGTTACCTTCCGAAAAAGTTGGTGTGATCTTGAATCCGCAAAACAAACA 4240
 QY 2439 CCGCTGTAGCGGTGTTTGTGTTTGTGCAAGCAGCAGATTACGCGCAAAAAAGAGAT 4300
 DB 4301 CCGCTGTAGCGGTGTTTGTGTTTGTGCAAGCAGCAGATTACGCGCAAAAAAGAGAT 4300
 QY 2499 CTCAAGAAAGTCTTTGATCTTTTCTTCTACGCGGTCTGACGCTAGAGAAAGAACTAC 4360
 DB 4361 CTCAAGAAAGTCTTTGATCTTTTCTTCTACGCGGTCTGACGCTAGAGAAAGAACTAC 4360
 QY 2559 GTTAAAGGATTTGGTGTATGATGATTAACAAAAAGATCTTCACTAGATCTTTAAATT 4420
 DB 4421 GTTAAAGGATTTGGTGTATGATGATTAACAAAAAGATCTTCACTAGATCTTTAAATT 4420
 QY 2619 AAAAATGAATTTTAAATCAATCTAAAGTATATAGAGTAAACTGTGCTACAGTTACC 4480
 DB 4481 AAAAATGAATTTTAAATCAATCTAAAGTATATAGAGTAAACTGTGCTACAGTTACC 4480
 QY 2679 AATGCTTAATCAGTGAAGGACCACTATCTCAGGATCTGTCTAATTTGTTCACTAGATTG 4540
 DB 4541 AATGCTTAATCAGTGAAGGACCACTATCTCAGGATCTGTCTAATTTGTTCACTAGATTG 4540
 QY 2739 CCGTACTCCCGTCTGTATGATTAACATCAGATACGAGGCTTCAATTTGTTCACTAGATTG 4600
 DB 4601 CCGTACTCCCGTCTGTATGATTAACATCAGATACGAGGCTTCAATTTGTTCACTAGATTG 4600
 QY 2799 CTGCAATGATACCGGAGACCCAGCTCAGCGGTCTCAGATTTATCAGCAATTAACAGC 4660
 DB 4661 CTGCAATGATACCGGAGACCCAGCTCAGCGGTCTCAGATTTATCAGCAATTAACAGC 4660
 QY 2859 CAGCCGGAAGGCGGAGACCGCAAGAGTGTCTGCAACTTATCCGCTCAGCTTCACTAGATTG 4720
 DB 4721 CAGCCGGAAGGCGGAGACCGCAAGAGTGTCTGCAACTTATCCGCTCAGCTTCACTAGATTG 4720
 QY 2919 TTAATTTGTGCGGGAAGTAGAGTAAAGTGTGCGCAATTAAGTTTGGCAACGTTG 4780
 DB 4781 TTAATTTGTGCGGGAAGTAGAGTAAAGTGTGCGCAATTAAGTTTGGCAACGTTG 4780
 QY 2979 TTGCAATTTGTGCAAGGATGTGTGTCAAGCTGCTGCTGTTGTTGTTGCTTCACTAGATTG 4840
 DB 4841 TTGCAATTTGTGCAAGGATGTGTGTCAAGCTGCTGCTGTTGTTGTTGCTTCACTAGATTG 4840
 QY 3039 CCGGTTCCCAACGATCAAGGAGTTTACATGATCCCAATGTTGTGCAAAAAAGCGTTA 4900
 DB 4901 CCGGTTCCCAACGATCAAGGAGTTTACATGATCCCAATGTTGTGCAAAAAAGCGTTA 4900
 QY 3099 GTCCTTGGGTCTCGATCTTGTGCAAGTAAAGTTGCGCAGTATTAATCACTAGATTG 4960
 DB 4961 GTCCTTGGGTCTCGATCTTGTGCAAGTAAAGTTGCGCAGTATTAATCACTAGATTG 4960

QY 3159 TTAATGAGCACTGATTAATCTCTTAATGATGATGATGATGATGATGATGATGATGAT 3218
 DB 5021 TTAATGAGCACTGATTAATCTCTTAATGATGATGATGATGATGATGATGATGATGAT 3218
 QY 3219 CTGTGATGATCTCAACCAAGTATCTTGAATGATGATGATGATGATGATGATGATGATGAT 5080
 DB 5081 CTGTGATGATCTCAACCAAGTATCTTGAATGATGATGATGATGATGATGATGATGATGAT 5080
 QY 3279 GCGCGGCTCAATACGGAATTAATACGCGCAGATGATGATGATGATGATGATGATGATGAT 5140
 DB 5141 GCGCGGCTCAATACGGAATTAATACGCGCAGATGATGATGATGATGATGATGATGATGAT 5140
 QY 3339 TTGAAAAAGTTCTCGGCGGAAAACTTCAAGATCTTACCGTGTGATGATGATGATGATGAT 5200
 DB 5201 TTGAAAAAGTTCTCGGCGGAAAACTTCAAGATCTTACCGTGTGATGATGATGATGATGAT 5200
 QY 3399 CGATGTAACCACTCTGTCACCACTGATCTTCAAGATCTTCACTTCACTTCACTTCACTTCA 5260
 DB 5261 CGATGTAACCACTCTGTCACCACTGATCTTCAAGATCTTCACTTCACTTCACTTCACTTCA 5260
 QY 3459 CTGGGTGAGCAAAAAAGGAAAGCAAAATGCGGAAAAAGGAAATAGGCGCAACGGA 5320
 DB 5321 CTGGGTGAGCAAAAAAGGAAAGCAAAATGCGGAAAAAGGAAATAGGCGCGGCTT 5320
 QY 3519 AATGTAATGATCTATCTTCTTCTTCAATTAATTAAGATTAATTAAGATTAATTAAGATTAAT 5380
 DB 5381 AATGTAATGATCTATCTTCTTCTTCAATTAATTAAGATTAATTAAGATTAATTAAGATTAAT 5380
 QY 3579 GTCTATGAGCGGATATATTTGATGATTTTGAATTAATTAATTAATTAATTAATTAATTAAT 5440
 DB 5441 GTCTATGAGCGGATATATTTGATGATTTTGAATTAATTAATTAATTAATTAATTAATTAAT 5440
 QY 3639 GCACATTTCCCGGAAAGTCCAC 3662
 DB 5501 GCACATTTCCCGGAAAGTCCAC 5524

RESULT 12
 US-09-573-322-20
 / Sequence 20, Application US/09573322
 / Patent No. 6531289
 / GENERAL INFORMATION:
 / APPLICANT: Bradley, John D.
 / APPLICANT: Thompson, Craig M.
 / APPLICANT: Moore, Jeffrey B.
 / APPLICANT: Mobbe, C. Richard
 / APPLICANT: Bailey, David A.
 / TITLE OF INVENTION: Regulated Gene Expression in Yeast and
 / FILE REFERENCE: 0342/ID469-US4
 / CURRENT APPLICATION NUMBER: US/09/573,322
 / PRIOR FILING DATE: 2000-05-18
 / PRIOR APPLICATION NUMBER: 09/404,066
 / PRIOR FILING DATE: 1999-09-23
 / PRIOR APPLICATION NUMBER: 09/138,024
 / PRIOR FILING DATE: 1998-08-21
 / PRIOR APPLICATION NUMBER: 60/056,719
 / NUMBER OF SEQ ID NOS: 30
 / SOFTWARE: FASTSEQ for Windows Version 3.0
 / SEQ ID NO 20
 / LENGTH: 7102
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: ZM195 plasmid
 US-09-573-322-20

Query Match 59.3%; Score 2172.8; DB 4; Length 7102;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1479 GAGTTACCCGCGGCGGCGCTTGGCTAATCATGTCATAGCTGTTTCTGTGTGAATTT 1538

Db	3341	GAGTGCACCGGATGCAAGCTTGGGGTATCATGCTCTAGTGTCTCTGTGTAATTT	3400
OY	1539	GTTATCCGCTCACAATTCACACAATACGAGCGCGAAGCTAAAGTGTAAAGCTGGG	1598
Db	3401	GTTATCCGCTCACAATTTCCACACAATACGAGCGCGAAGCTAAAGTGTAAAGCTGGG	3460
OY	1599	GTGCTTAATGATGAGCTAACTCACTTAATTGCGTTCGCTCACTGCGCGCTTCCAGT	1658
Db	3461	GTGCTTAATGATGAGCTAACTCACTTAATTGCGTTCGCTCACTGCGCGCTTCCAGT	3520
OY	1659	CGGAAACCTGTCCGTCCAGCTGCATTAATGATGGCCAAACGCGCGGGAGAGCGGTT	1718
Db	3521	CGGAAACCTGTCCGTCCAGCTGCATTAATGATGGCCAAACGCGCGGGAGAGCGGTT	3580
OY	1719	TGCGTAATGGAGCGCTCTCCGCTTCTGAGCTCACTGACGTGGCGCTGGTCCGTGGC	1778
Db	3581	TGCGTAATGGAGCGCTCTCCGCTTCTGAGCTCACTGACGTGGCGCTGGTCCGTGGC	3640
OY	1779	TGCGGCGAGCGGTATCAGCTCATCTCAAAGCGGTAAATACGCTTAATCCACAGATCAGGG	1838
Db	3641	TGCGGCGAGCGGTATCAGCTCATCTCAAAGCGGTAAATACGCTTAATCCACAGATCAGGG	3700
OY	1839	ATAACGCGAAGAAACAATGTAGCAAAAAGCCAGCAAAAGGCCAGAACCGTAAAAAGG	1898
Db	3701	ATAACGCGAAGAAACAATGTAGCAAAAAGCCAGCAAAAGGCCAGAACCGTAAAAAGG	3760
OY	1899	CGCGCTTCTGAGCGTTTTCATAGGCTCCGCGCCCCCGTAGAGCATACAAAAATCGAC	1958
Db	3761	CGCGCTTCTGAGCGTTTTCATAGGCTCCGCGCCCCCGTAGAGCATACAAAAATCGAC	3820
OY	1959	GCTCAAGTCAGAGGTGGCGAAACCCGACAGACTTAAGAATACAGGCGTTCCCGCTG	2018
Db	3821	GCTCAAGTCAGAGGTGGCGAAACCCGACAGACTTAAGAATACAGGCGTTCCCGCTG	3880
OY	2019	GAACTCCCTGTGCGCTCTCTCTGTTCCGACCTTCGCGCTTAACGGATACCTGTCCGCT	2078
Db	3881	GAACTCCCTGTGCGCTCTCTCTGTTCCGACCTTCGCGCTTAACGGATACCTGTCCGCT	3940
OY	2079	TTCTCCCTTGGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGAAGATATCTCACTTCCG	2138
Db	3941	TTCTCCCTTGGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGAAGATATCTCACTTCCG	4000
OY	2139	TGTAGTGTGTTGCTCCAAAGCTGGGCTGTGACGAACCCCGCTTACGCCACCGCT	2198
Db	4001	TGTAGTGTGTTGCTCCAAAGCTGGGCTGTGTGACGAACCCCGCTTACGCCACCGCT	4060
OY	2199	GCGCTTAATCCGGTATCATATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTAATCGCAC	2258
Db	4061	GCGCTTAATCCGGTATCATATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTAATCGCAC	4120
OY	2259	TGGCAGCAGCCACTGTGTAAACAGATATACAGACGAGGTATGTAGCGGTCTCAAGAT	2318
Db	4121	TGGCAGCAGCCACTGTGTAAACAGATATACAGACGAGGTATGTAGCGGTCTCAAGAT	4180
OY	2319	TCTTGAATGGTGGGCTTAAGTACGAGCTCACTAAGAGACAGTATTTGGTATCTGCGCTC	2378
Db	4181	TCTTGAATGGTGGGCTTAAGTACGAGCTCACTAAGAGACAGTATTTGGTATCTGCGCTC	4240
OY	2379	TGCTGAAGCCAGTTACTCTTGGAAAAAAGATTGGTAGCTCTTGATTCGGCAAAACA	2438
Db	4241	TGCTGAAGCCAGTTACTCTTGGAAAAAAGATTGGTAGCTCTTGATTCGGCAAAACA	4300
OY	2439	CGCGCTGAGGGGTGGTTTTTTTGTTCGACAGCAGACATTAACGCGCAAAAAAGAT	2498
Db	4301	CGCGCTGAGGGGTGGTTTTTTTGTTCGACAGCAGACATTAACGCGCAAAAAAGAT	4360
OY	2499	CTCAAGAAGATCCCTTGTGATCTTTTCTACAGGGGTCTGACGCTCAGTGGAAAGAAACTCAC	2558
Db	4361	CTCAAGAAGATCCCTTGTGATCTTTTCTACAGGGGTCTGACGCTCAGTGGAAAGAAACTCAC	4420
OY	2559	GTTAAGGATTTTGGTCAATGATTAATCAAAAAAGATCTTCACTTAATCTTTTAATTT	2618

Db	4421	GTAAAGGATTTTGGTCATGAGATTATCAAAAAGATCTTCAACCTAGATCCTTTAAATT	4480
Qy	2619	AAAAATGAAGTTTAAATCAATCTTAAGATATATAGATAAATCTTGCTGACAGTTACC	2678
Db	4481	AAAAATGAAGTTTAAATCAATCTTAAGATATATAGATAAATCTTGCTGACAGTTACC	4540
Qy	2679	AATGCTTAATCAGTGAAGGACCTATCTCAAGGATCTGTCTAATTTTGTTTCATCATATGTTG	2738
Db	4541	AATGCTTAATCAGTGAAGGACCTATCTCAAGGATCTGTCTAATTTTGTTTCATCATATGTTG	4600
Qy	2739	CCTGACTCCCCGTGTGTAGATTACTACGATACGGGAGGGCTTACATCTGCCCCAGTG	2798
Db	4601	CCTGACTCCCCGTGTGTAGATTACTACGATACGGGAGGGCTTACATCTGCCCCAGTG	4660
Qy	2799	CTGCATATGATCCCGGAGACCCACGCTCACCGGCTCCAGATTATTCAGCAATAAACGAC	2858
Db	4661	CTGCATATGATCCCGGAGACCCACGCTCACCGGCTCCAGATTATTCAGCAATAAACGAC	4720
Qy	2859	CAGCCGAGAGGAGCCGAGCGCAGAAAGTGGTCTGCACTTTATCCGCTCCATCCAGTCTA	2918
Db	4721	CAGCCGAGAGGAGCCGAGCGCAGAAAGTGGTCTGCACTTTATCCGCTCCATCCAGTCTA	4780
Qy	2919	TTAATTTGTTCCGGGAAAGCTTAGAGTAAGTGTCCGCAATTAAATGTTTGGCAAGCTTG	2978
Db	4781	TTAATTTGTTCCGGGAAAGCTTAGAGTAAGTGTCCGCAATTAAATGTTTGGCAAGCTTG	4840
Qy	2979	TTGCAATTTGCTACAGGCGCATGTGGTGTACAGCTGTGTGTTGGATAGGCTTCAATCAGCT	3038
Db	4841	TTGCAATTTGCTACAGGCGCATGTGGTGTACAGCTGTGTGTTGGATAGGCTTCAATCAGCT	4900
Qy	3039	CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGTGCAAAAAAGCGTTA	3098
Db	4901	CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGTGCAAAAAAGCGTTA	4960
Qy	3099	GCTCTCTTGGTCTCTCCGATCGTTTCAAGATTAAGTTGGCCGCAAGTGTATCACTACATGG	3158
Db	4961	GCTCTCTTGGTCTCTCCGATCGTTTCAAGATTAAGTTGGCCGCAAGTGTATCACTACATGG	5020
Qy	3159	TTATGGGACGACCTGCAATTAATCTCTTACGTGCAATGCCATCCGTAAAGATGCTTTCTGTGA	3218
Db	5021	TTATGGGACGACCTGCAATTAATCTCTTACGTGCAATGCCATCCGTAAAGATGCTTTCTGTGA	5080
Qy	3219	CTGGTGAATACTCAACCAAGTCATTCTGAGAAATAGTATAGCGGCGACCGAGTTGCTCTT	3278
Db	5081	CTGGTGAATACTCAACCAAGTCATTCTGAGAAATAGTATAGCGGCGACCGAGTTGCTCTT	5140
Qy	3279	GCCCCGGGCTCAATACGGGATATATCCGCGCACATACGACGAACTTTAAAGTGTCTATCA	3338
Db	5141	GCCCCGGGCTCAATACGGGATATATCCGCGCACATACGACGAACTTTAAAGTGTCTATCA	5200
Qy	3339	TTGAAAAACGTTCTTGGGGGCGAAACCTCTCAAGGATCTTAACGCTGTAGAGATCCAGTT	3398
Db	5201	TTGAAAAACGTTCTTGGGGGCGAAACCTCTCAAGGATCTTAACGCTGTAGAGATCCAGTT	5260
Qy	3399	CGATGTAAACCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACCTTACACGCGTTT	3458
Db	5261	CGATGTAAACCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACCTTACACGCGTTT	5320
Qy	3459	CTGGGTGAGCAAAAACGAAAGGCAAAATGCGCAAAAAAGGAAATAGGGCGACACGGA	3518
Db	5321	CTGGGTGAGCAAAAACGAAAGGCAAAATGCGCAAAAAAGGAAATAGGGCGACACGGA	5380
Qy	3519	AATGTTGAATATCTATCTCTCTCTTTTCAATATTAATGAAACATTTTATCAGGGTTATT	3578
Db	5381	AATGTTGAATATCTATCTCTCTCTTTTCAATATTAATGAAACATTTTATCAGGGTTATT	5440
Qy	3579	GTCTCATAGAGCGATACATATTTGAATGTTTGAATAAAATTAACAATAGGGGTTCCGC	3638
Db	5441	GTCTCATAGAGCGATACATATTTGAATGTTTGAATAAAATTAACAATAGGGGTTCCGC	5500
Qy	3639	GCACATTTTCCCGAAAAAGTGCAC 3662	
Db	5501	GCACATTTTCCCGAAAAAGTGCAC 5524	

Db 5256 TTATGGACGACTGATTAATTTCTTACTGTCATCCGTAAGATGCTTTTCTGCA 5315
Qy 3219 CTGGTGAAGTACTCAACCAAGTCATTTCTGAGAAATAGTATGCGGACCGAGTTCTCTT 3278
Db 5316 CTGGTGAAGTACTCAACCAAGTCATTTCTGAGAAATAGTATGCGGACCGAGTTCTCTT 5375
Qy 3279 GCGCGGCGTCAATAGCGGATTAATACCGGCGCATAGAGAACTTTAAAGTCTCATCA 3338
Db 5376 GCGCGGCGTCAATAGCGGATTAATACCGGCGCATAGAGAACTTTAAAGTCTCATCA 5435
Qy 3339 TTGGAACAGTCTTCTGCGGCGCAAACTCTCAAGAGCTTACCGCTGTGAGATCAAGTT 3398
Db 5436 TTGGAACAGTCTTCTGCGGCGCAAACTCTCAAGAGCTTACCGCTGTGAGATCAAGTT 5495
Qy 3399 CGATGTAAACCACTGTCGACCCAACTGATCTTCAGCATCTTTTACCTTCAACAGCGTTT 3458
Db 5496 CGATGTAAACCACTGTCGACCCAACTGATCTTCAGCATCTTTTACCTTCAACAGCGTTT 5555
Qy 3459 CTGGGTGACAAAAACAGAAAGCAAAATGCGCAAAAAAGGAAATTAAGGCGACACGGA 3518
Db 5556 CTGGGTGACAAAAACAGAAAGCAAAATGCGCAAAAAAGGAAATTAAGGCGACACGGA 5615
Qy 3519 AATGTTGAATCTCATCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTAT 3578
Db 5616 AATGTTGAATCTCATCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTAT 5675
Qy 3579 GTCTCATAGCGGATACATATTTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGC 3638
Db 5676 GTCTCATAGCGGATACATATTTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGC 5735
Qy 3639 GCACATTTCCCGAAAGTGCCAC 3662
Db 5736 GCACATTTCCCGAAAGTGCCAC 5759

RESULT 14
US-09-404-066-21
Sequence 21, Application US/09404066
Patent No. 6365409
GENERAL INFORMATION:
APPLICANT: Bradley, John D.
APPLICANT: Thompson, Craig M.
APPLICANT: Moore, Jeffrey B.
APPLICANT: Mobbe, C. Richard
APPLICANT: Healy, Judith M.
FILE REFERENCE: 0342/1D469U51
CURRENT APPLICATION NUMBER: US/09/404,066
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/138,024
PRIOR FILING DATE: 1998-08-21
PRIOR APPLICATION NUMBER: 60/056,719
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq For Windows Version 3.0
SEQ ID NO 21
LENGTH: 7333
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Plasmid pZM197
US-09-404-066-21

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1479 GAGTTACCGGCGCGGCGGCTTGGCGTATCATGTCTATCTGTTCTGTGAAAT 1538
Db 3576 GAGTCAGCGGCGGCGGCGGCTTGGCGTATCATGTCTATCTGTTCTGTGAAAT 3635

Qy 1539 GTTATCCGCTCAATTCACACAACATAGAGCGGAGAGCTAAAGTAAAGCTGGG 1598
Db 3636 GTTATCCGCTCAATTCACACAACATAGAGCGGAGAGCTAAAGTAAAGCTGGG 3695
Qy 1599 GTGCTTAATGAGTGAAGTCACTCAATTAATGCGTTGCGCTCACTGCCCTTTCAGT 1658
Db 3696 GTGCTTAATGAGTGAAGTCACTCAATTAATGCGTTGCGCTCACTGCCCTTTCAGT 3755
Qy 1659 CGGGAACCTGTCTGCGCAAGTCAATTAATATGCGCAACCGCGGAGAGAGCGGTT 1718
Db 3756 CGGGAACCTGTCTGCGCAAGTCAATTAATATGCGCAACCGCGGAGAGAGCGGTT 3815
Qy 1719 TGCGTATGAGGCGCTTTCGCTTCCTGCTCACTGATCGCTGCGCTGCTGCTGCTGCTG 1778
Db 3816 TGCGTATGAGGCGCTTTCGCTTCCTGCTCACTGATCGCTGCGCTGCTGCTGCTGCTG 3875
Qy 1779 TGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTATATAGGTTATCCACAAATCAGGG 1838
Db 3876 TGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTATATAGGTTATCCACAAATCAGGG 3935
Qy 1839 ATAAACGAGAAAGAACATGTGACCAAAAGGCGCAAAAGGCGCAAGAACCGTAAAGG 1898
Db 3936 ATAAACGAGAAAGAACATGTGACCAAAAGGCGCAAAAGGCGCAAGAACCGTAAAGG 3995
Qy 1899 CCGGCTGCTGCGCTTTTCCATAGGCTCGCGCCCCCGGAGAGCATCAAAAAATCGAC 1958
Db 3996 CCGGCTGCTGCGCTTTTCCATAGGCTCGCGCCCCCGGAGAGCATCAAAAAATCGAC 4055
Qy 1959 GCTCAAGTCAAGGTGCGGAACCGGACAGACTATAAGATATACAGGCTTTCCCGCTG 2018
Db 4056 GCTCAAGTCAAGGTGCGGAACCGGACAGACTATAAGATATACAGGCTTTCCCGCTG 4115
Qy 2019 GAAGCTCCCTGTCGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCT 2078
Db 4116 GAAGCTCCCTGTCGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCT 4175
Qy 2079 TTCTCCCTTGGGGAAGCGTGGCGCTTCTCATCTCACTGATAGATCTCAAGTTCCG 2138
Db 4176 TTCTCCCTTGGGGAAGCGTGGCGCTTCTCATCTCACTGATAGATCTCAAGTTCCG 4235
Qy 2139 TGTAGTGTGTCGCTCAAGCTGAGCTGTGACAGAAACCCCGTTACGCCACGCT 2198
Db 4236 TGTAGTGTGTCGCTCAAGCTGAGCTGTGACAGAAACCCCGTTACGCCACGCT 4295
Qy 2199 GCGCTTATCCGTAATCATATGCTTGAAGTCAACCCGTTAAGACAGACTTATCGCAC 2258
Db 4296 GCGCTTATCCGTAATCATATGCTTGAAGTCAACCCGTTAAGACAGACTTATCGCAC 4355
Qy 2259 TGGCAGACGCACTGTGTAACAGATTAAGCAGAGGATATGAGGCGGTCTACAGGT 2318
Db 4356 TGGCAGACGCACTGTGTAACAGATTAAGCAGAGGATATGAGGCGGTCTACAGGT 4415
Qy 2319 TCTTGAAGTGTGCTTAACCTAAGGCTAAGCTAAGGACAGTATTTGGTATCTGCGCTC 2378
Db 4416 TCTTGAAGTGTGCTTAACCTAAGGCTAAGCTAAGGACAGTATTTGGTATCTGCGCTC 4475
Qy 2439 TGCTGAAGCCAGTTACCTTGGAAAAAGATTGAGTACTTGAATCCGCAACAAACCA 2438
Db 4476 TGCTGAAGCCAGTTACCTTGGAAAAAGATTGAGTACTTGAATCCGCAACAAACCA 4535
Qy 2439 CCGCTGTGAGCGGCTTTTGTGTTGCAAGCAGAGATTAACCGCAAAAAAGAT 2498
Db 4536 CCGCTGTGAGCGGCTTTTGTGTTGCAAGCAGAGATTAACCGCAAAAAAGAT 4595
Qy 2499 CTCAAGAAAGATCTTTGATCTTTCTGCGGCGTCAAGCTCAGTGGACGAAGAACTCAC 2558
Db 4596 CTCAAGAAAGATCTTTGATCTTTCTGCGGCGTCAAGCTCAGTGGACGAAGAACTCAC 4655
Qy 2559 GTTAAGGATTTTGTGATGATTAATCAAAAGATCTTCACTAGATCTTTTAAT 2618
Db 4656 GTTAAGGATTTTGTGATGATTAATCAAAAGATCTTCACTAGATCTTTTAAT 4715
Qy 2619 AAAATAGATTTTAATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTACC 2678

4716 AAAAAAGAGTTTAAATCAATCTAAAGATATATAGAACTTGGATCAAGCTTACC 4775
QY 2679 AATGCTTAATCACTAGAGGACCTATCTCAGCATCTGCTATTTCTTCTATCCATAGTTG 2738
Db 4776 AATGCTTAATCACTAGAGGACCTATCTCAGCATCTGCTATTTCTTCTATCCATAGTTG 2738
QY 2739 CCTGATCCCGGTCGGTAGATTAATCAAGATACGGAGAGGCTTACACTTGGCCCCAGTG 4835
Db 4836 CCTGATCCCGGTCGGTAGATTAATCAAGATACGGAGAGGCTTACACTTGGCCCCAGTG 4835
QY 2799 CTGCATGATACCGGAGAACCCACGCTCAGCGGCTCCAGATTTATCGCAATTAACACG 2858
Db 4896 CTGCATGATACCGGAGAACCCACGCTCAGCGGCTCCAGATTTATCGCAATTAACACG 4955
QY 2859 CAGCGGAGAGGCGGAGGAGGAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA 2918
Db 4956 CAGCGGAGAGGCGGAGGAGGAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTA 2918
QY 2919 TTAATTTGTCGGGAAAGCTAGATAGTATGTCGACAGTTATAGTTTGGCCAACTTG 5015
Db 5016 TTAATTTGTCGGGAAAGCTAGATAGTATGTCGACAGTTATAGTTTGGCCAACTTG 5015
QY 2979 TTGCAATTTGCTACAGGAGTATGTTGTTGACGCTGCTGTTGGTATGCTTCAATCAGCT 3038
Db 5076 TTGCAATTTGCTACAGGAGTATGTTGTTGACGCTGCTGTTGGTATGCTTCAATCAGCT 3038
QY 3039 CCGGTTCCCAACGATCAAGGCGAGTACATGATCCCGGCTGTTGGCAAAAAGGCGTTA 3098
Db 5136 CCGGTTCCCAACGATCAAGGCGAGTACATGATCCCGGCTGTTGGCAAAAAGGCGTTA 3098
QY 3099 GCTCCTTCGGTCTCCGATCTGTTGTGAGATAGTATGTCGAGTATGTTTATCACTCATG 3158
Db 5196 GCTCCTTCGGTCTCCGATCTGTTGTGAGATAGTATGTCGAGTATGTTTATCACTCATG 3158
QY 3159 TTAATGACAGCACTGATTAATCTCTTAATCTGATCCGCTGTTGAGATGCTTTTCTGTA 3218
Db 5256 TTAATGACAGCACTGATTAATCTCTTAATCTGATCCGCTGTTGAGATGCTTTTCTGTA 3218
QY 3219 CTGTGATGATCAACCAAGTATCTGATTAATGATGATGTCGAGACCGAGTTGCTT 5315
Db 5316 CTGTGATGATCAACCAAGTATCTGATTAATGATGATGTCGAGACCGAGTTGCTT 5315
QY 3279 GCGCGGCGTCAATACGAGGATTAATACCGGCGACATACAGAACTTTAAAGTCTCATCA 3338
Db 5376 GCGCGGCGTCAATACGAGGATTAATACCGGCGACATACAGAACTTTAAAGTCTCATCA 3338
QY 3339 TTGAAAAAGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCAAGTT 3398
Db 5436 TTGAAAAAGTTCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCAAGTT 3398
QY 3399 CGATGTAAACCACTCGTGAACCCCACTGATCTTACGATCTTTTACTTCAACAGGCTT 3458
Db 5496 CGATGTAAACCACTCGTGAACCCCACTGATCTTACGATCTTTTACTTCAACAGGCTT 3458
QY 3459 CTGGGTAGGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5555
Db 5556 CTGGGTAGGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5555
QY 3519 AATGTTGATATCTCATCTCTCTTTTCAATATTAATTAAGAGCTTTATCAAGGTTAT 3578
Db 5616 AATGTTGATATCTCATCTCTCTTTTCAATATTAATTAAGAGCTTTATCAAGGTTAT 3578
QY 3579 GTCTCATGAGCGGATCATATTTGAATGTTAGAAAAATTAACAATAGAGGTTCCGC 3638
Db 5676 GTCTCATGAGCGGATCATATTTGAATGTTAGAAAAATTAACAATAGAGGTTCCGC 3638
QY 3639 GCAATTTTCCCGAAAAAGTGCAC 3662
Db 5736 GCAATTTTCCCGAAAAAGTGCAC 5759

RESULT 15

US-09-573-322-21
/ Sequence 21, Application US/09573322
/ Patent No. 6531289
/ GENERAL INFORMATION:
/ APPLICANT: Bradley, John D.
/ APPLICANT: Thompson, Craig M.
/ APPLICANT: Moore, Jeffrey B.
/ APPLICANT: Mobbe, C. Richard
/ APPLICANT: Bailey, David A.
/ TITLE OF INVENTION: Regulated Gene Expression in Yeast and
/ FILE REFERENCE: Methods of Use
/ CURRENT APPLICATION NUMBER: US/09/573,322
/ PRIOR FILING DATE: 2000-05-18
/ PRIOR APPLICATION NUMBER: 09/404,066
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: 09/138,024
/ PRIOR FILING DATE: 1998-08-21
/ PRIOR APPLICATION NUMBER: 60/056,719
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 21
/ LENGTH: 7333
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: ZM197 plasmid
US-09-573-322-21

Query Match 59.3%; Score 2172.8; DB 4; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1479 GAGTTACCCGCGGCGCGCTTGGCGTATCATGTATCATAGCTGTTCTGTGTAAATT 1538
Db 3576 GAGTGAACCGGACGACGAAAGCTGGCGTATCATGTATCATAGCTGTTCTGTGTAAATT 1538
QY 1539 GTTATCCGCTCAACAATTCACACACATACAGAGCGGAAAGATTAAGCTGG 1598
Db 3636 GTTATCCGCTCAACAATTCACACACATACAGAGCGGAAAGATTAAGCTGG 1598
QY 1599 GTGCTTAATGATGAGCTAACTCAATTAATGCTTGGCTGCTGCTGCTTCCAGT 1658
Db 3696 GTGCTTAATGATGAGCTAACTCAATTAATGCTTGGCTGCTGCTGCTTCCAGT 1658
QY 1659 CCGGAAACCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3755
Db 3756 CCGGAAACCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3755
QY 1719 TGGGTATTGGGCGCTCTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3815
Db 3816 TGGGTATTGGGCGCTCTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3815
QY 1779 TGGGCGAGCGGTATACAGTCACTCAAAAGGCGGTAAATCGGTTATCCAGAAAT 1838
Db 3876 TGGGCGAGCGGTATACAGTCACTCAAAAGGCGGTAAATCGGTTATCCAGAAAT 1838
QY 1839 ATTAAGCGAGAAAGCAATGTAGCAAAAGGCGCAAAAGGCGCAAGAACCTGTAAGAG 1898
Db 3936 ATTAAGCGAGAAAGCAATGTAGCAAAAGGCGCAAAAGGCGCAAGAACCTGTAAGAG 1898
QY 1899 CCGCGTCTGCTGCTGCTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGAC 1958
Db 3996 CCGCGTCTGCTGCTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGAC 1958
QY 1959 GCTCAAGTCAGAGGTGCGAAACCGGACAGGATTAAGGATTAACGAGGCTTCCCGCTG 4055
Db 4056 GCTCAAGTCAGAGGTGCGAAACCGGACAGGATTAAGGATTAACGAGGCTTCCCGCTG 4055
QY 2019 GAAGCTCCCTGAGGCTCTCCCTGTTCCGACCTGCGGTTACCGGATACCTGTCCGCT 2078
Db 4116 GAAGCTCCCTGAGGCTCTCCCTGTTCCGACCTGCGGTTACCGGATACCTGTCCGCT 4115

QY 2079 TTCTCCCTTGGGAGCGGTGGCGCTTTCATAGCTCAAGCTAGATCTCACTTGGG 2138
 Db 4176 TTCTCCCTTGGGAGCGGTGGCGCTTTCATAGCTCAAGCTAGATCTCACTTGGG 4235
 QY 2139 TGTAGTGTGTGGCTCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCCGACT 2198
 Db 4236 TGTAGTGTGTGGCTCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCCGACT 4295
 QY 2199 GCGCTTATCCGGTACTATGTCTTGAATCCACCCGGTAAGACAGACTTATGCGAC 2258
 Db 4296 GCGCTTATCCGGTACTATGTCTTGAATCCACCCGGTAAGACAGACTTATGCGAC 4355
 QY 2259 TGGAGAGCGCACTGTGTAAAGATTAAGACAGAGATTAAGAGCGGTCTACAGAT 2318
 Db 4356 TGGAGAGCGCACTGTGTAAAGATTAAGACAGAGATTAAGAGCGGTCTACAGAT 4415
 QY 2319 TCTTGAAGTGTGGCTTAAGCTAAGGCTACAGTAAGAGACAGTATTTGTATCTGCGCTC 2378
 Db 4416 TCTTGAAGTGTGGCTTAAGCTAAGGCTAAGTAAGAGACAGTATTTGTATCTGCGCTC 4475
 QY 2379 TGTGAAAGCGCTTACCTTCCGAAAAGAGTTGGTACTTGTATCCGGCAAAACCA 2438
 Db 4476 TGTGAAAGCGCTTACCTTCCGAAAAGAGTTGGTACTTGTATCCGGCAAAACCA 4535
 QY 2439 CCGCTGGTAGCGGTGCTTTTGTGTTGCAAGCAGCAATTAACGGCGCAAAAAGAGAT 2498
 Db 4536 CCGCTGGTAGCGGTGCTTTTGTGTTGCAAGCAGCAATTAACGGCGCAAAAAGAGAT 4595
 QY 2499 CTGAAGAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAATCAC 2558
 Db 4596 CTGAAGAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAATCAC 4655
 QY 2559 GTTAAAGGATTTTGGTCAATGATTAATCAAAAAGAGATTTGACCTGATCTTTAAAT 2618
 Db 4656 GTTAAAGGATTTTGGTCAATGATTAATCAAAAAGAGATTTGACCTGATCTTTAAAT 4715
 QY 2619 AAAAATGAAGTTTAAATCAATCTAAATATTAATGAATTAATTTGCTGACAGTTAC 2678
 Db 4716 AAAAATGAAGTTTAAATCAATCTAAATATTAATGAATTAATTTGCTGACAGTTAC 4775
 QY 2679 AATGCTTAAATGATGAGGCACTATCTCAGGATCTGTCTATTTGCTTCAATCCATAGTTG 2738
 Db 4776 AATGCTTAAATGATGAGGCACTATCTCAGGATCTGTCTATTTGCTTCAATCCATAGTTG 4835
 QY 2739 CTTGATCTCCCGTGTGTAGTAACTACGATACGGAGGCTTACCATCTGCCCCAGTG 2798
 Db 4836 CTTGATCTCCCGTGTGTAGTAACTACGATACGGAGGCTTACCATCTGCCCCAGTG 4895
 QY 2799 CTGCAATGATACCGCGAGACCCACGCTCAACGGCTCCAGATTTATCAGCAATAAACGAC 2858
 Db 4896 CTGCAATGATACCGCGAGACCCACGCTCAACGGCTCCAGATTTATCAGCAATAAACGAC 4955
 QY 2859 CAGCCGGAAGGGCCGAGCGCAGAAAGTGTCTGCAACTTATCCGCCCTCCATCCAGTCTA 2918
 Db 4956 CAGCCGGAAGGGCCGAGCGCAGAAAGTGTCTGCAACTTATCCGCCCTCCATCCAGTCTA 5015
 QY 2919 TTAATTTGTCGGGAAAGCTAGAGTAAGTAAGTTGCGCAATTATGTTGCGCAACGTTG 2978
 Db 5016 TTAATTTGTCGGGAAAGCTAGAGTAAGTAAGTTGCGCAATTATGTTGCGCAACGTTG 5075
 QY 2979 TTGCAATTTGCTACAGGCACTGCTGTGTCAACGCTGCTTGTGTATGCTTCAATTCAGCT 3038
 Db 5076 TTGCAATTTGCTACAGGCACTGCTGTGTGTCAACGCTGCTTGTGTATGCTTCAATTCAGCT 5135
 QY 3039 CCGGTTCCCAACGATCAAGCGAGTTCATGATCCCCCATTTGTGCAAAAAGCGGTTA 3098
 Db 5136 CCGGTTCCCAACGATCAAGCGAGTTCATGATCCCCCATTTGTGCAAAAAGCGGTTA 5195
 QY 3099 GCTCTTTCGGTCTTCCGATCGTTGTCAAGATTAAGTTGCGCAGTGTATCACTCATGG 3158
 Db 5196 GCTCTTTCGGTCTTCCGATCGTTGTCAAGATTAAGTTGCGCAGTGTATCACTCATGG 5255

QY 3159 TTATGGAGCACTGCATTAATCTCTTACTGTGCATGCCATCCGTAAGATGCTTTCTGTGA 3218
 Db 5256 TTATGGAGCACTGCATTAATCTCTTACTGTGCATGCCATCCGTAAGATGCTTTCTGTGA 5315
 QY 3219 CTGTGAGTACTCAACCAAGTCAATTTGAGAATAGTATGAGCGGACCGAGTTGCTCTT 3278
 Db 5316 CTGTGAGTACTCAACCAAGTCAATTTGAGAATAGTATGAGCGGACCGAGTTGCTCTT 5375
 QY 3279 GCCCGGCGTCAATACGGGATTAATACCGGCGCAATAGCAGAACTTTAAAGTCTCATCA 3338
 Db 5376 GCCCGGCGTCAATACGGGATTAATACCGGCGCAATAGCAGAACTTTAAAGTCTCATCA 5435
 QY 3339 TTGAAAAACGTTCTTCCGGGGGAAAACCTCAAGATCTTACCGCTGTGAGATCCAGTT 3398
 Db 5436 TTGAAAAACGTTCTTCCGGGGGAAAACCTCAAGATCTTACCGCTGTGAGATCCAGTT 5495
 QY 3399 CGATGTAAACCACTCGTGACCCCACTGATCTTCAAGCATCTTTTACTTTTCAACGCGTTT 3458
 Db 5496 CGATGTAAACCACTCGTGACCCCACTGATCTTCAAGCATCTTTTACTTTTCAACGCGTTT 5555
 QY 3459 CTGGGTGACAAAAACAGAAAGCAAAATGCGCAAAAAGGAAATTAAGGCGACACCGA 3518
 Db 5556 CTGGGTGACAAAAACAGAAAGCAAAATGCGCAAAAAGGAAATTAAGGCGACACCGA 5615
 QY 3519 AATGTGAATCTCATCTCTTCTTTTCAATTAATGAAGATTTATCAGGTTAAT 3578
 Db 5616 AATGTGAATCTCATCTCTTCTTTTCAATTAATGAAGATTTATCAGGTTAAT 5675
 QY 3579 GTCTCATGAGCGGATTAATTTGAATGATTTTGAATAAATAAAGGAGTTCCGC 3638
 Db 5676 GTCTCATGAGCGGATTAATTTGAATGATTTTGAATAAATAAAGGAGTTCCGC 5735
 QY 3639 GCACATTTCCCGAAAAAGTGCAC 3662
 Db 5736 GCACATTTCCCGAAAAAGTGCAC 5759

Search completed: January 26, 2005, 12:21:01
 Job time : 362.259 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:27:35 ; Search time 3011.51 Seconds

(without alignments)
11102.026 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 1 ggcaccatgcgcgcacattt.....aaatgcagcagcgatcc 707

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hng:*
3: gb_in:*
4: gb_cm:*
5: gb_cv:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.4	82.8	20067	12 XXU02425	U02425 Cloning vec
2	585.4	82.8	20125	12 XXU02426	U02426 Cloning vec
3	585.4	82.8	42529	12 CVU39284	U039284 Cloning vec
4	585.4	82.8	42530	12 CVU39285	U039285 Cloning vec
5	585.4	82.8	42531	12 CVU39286	U039286 Cloning vec
6	585.4	82.8	42704	12 CVU37692	U037692 Cloning vec
7	585.4	82.8	48502	7 LAMCG	J02459 Bacterioph
8	574.6	81.3	5965	12 AF434923	AF434923 Expressio
9	574.6	81.3	6182	12 AF434925	AF434925 Expressio
10	574.6	81.3	6182	12 AF434926	AF434926 Expressio
11	574.6	81.3	6182	12 AF434927	AF434927 Expressio
12	574.6	81.3	6182	12 AF434931	AF434931 Expressio
13	574.6	81.3	6182	12 AF434932	AF434932 Expressio
14	574.6	81.3	7421	12 AF434924	AF434924 Expressio
15	574.6	81.3	7634	12 AF434928	AF434928 Expressio
16	574.6	81.3	7634	12 AF434929	AF434929 Expressio
17	574.6	81.3	7634	12 AF434930	AF434930 Expressio
18	574.6	81.3	7634	12 AF434933	AF434933 Expressio
19	574.6	81.3	7634	12 AF434934	AF434934 Expressio

20	566.2	80.1	14164	1 AE005330	AE005330 Escherich
21	566.2	80.1	26565	6 AX702505	AX702505 Sequence
22	566.2	80.1	46819	6 AR204176	AR204176 Sequence
23	566.2	80.1	305325	1 AE016765	AE016765 Escherich
24	564.6	79.9	46897	6 BD184772	BD184772 Nucleic a
25	564.6	79.9	222605	1 AP002555	AP002555 Escherich
26	561.4	79.4	22305	6 AR204355	AR204355 Sequence
27	561.4	79.4	91772	6 BD184765	BD184765 Nucleic a
28	561.4	79.4	286485	1 AP002557	AP002557 Escherich
29	407	57.6	46375	7 AF064539	AF064539 Bacteriop
30	395.8	56.0	64137	2 AC101160	AC101160 Mus muscu
31	395.8	56.0	179840	2 AC068458	AC068458 Homo sapi
32	395.8	56.0	179871	2 AC024602	AC024602 Homo sapi
33	395.8	56.0	194028	2 AC111077	AC111077 Homo sapi
34	395.8	56.0	206486	2 AC074218	AC074218 Mus muscu
35	362.2	51.2	228516	2 AC008594	AC008594 Homo sapi
36	343.2	48.5	68028	2 AC101473	AC101473 Mus muscu
37	329.8	46.6	66450	2 AC060821	AC060821 Homo sapi
38	320.4	45.3	63579	2 AC100018	AC100018 Mus muscu
39	320.2	45.3	63690	2 AC125419	AC125419 Homo sapi
40	231.8	32.8	29466	1 AE008818	AE008818 Salmonell
41	223	31.5	8134	7 PH2HEADTL	PH2HEADTL
42	223	31.5	13264	1 AE005368	AE005368 Escherich
43	223	31.5	13386	1 AE015100	AE015100 Shigella
44	223	31.5	34063	6 AR204200	AR204200 Sequence
45	223	31.5	44029	6 BD184774	BD184774 Nucleic a

ALIGNMENTS

RESULT 1
XXU02425 20067 bp DNA linear SYN 29-JAN-1997
LOCUS Cloning vector lambda EMBL3, left arm.
DEFINITION U02425
ACCESSION U02425
VERSION U02425.1 GI:413791
KEYWORDS
SOURCE
ORGANISM
Cloning vector lambda EMBL3
Cloning vector lambda EMBL3
artificial sequences; vectors.
REFERENCE
1 (bases 1 to 20067)
Frischauf, A.M., Lehrach, H., Poustka, A. and Murray, N.
Lambda replacement vectors carrying polylinker sequences
J. Mol. Biol. 170 (4), 827-842 (1983)
JOURNAL MEDLINE
PUBMED
6315951
REFERENCE
2 (bases 1 to 20067)
Kitts, P.A.
Clontechn Vectors On Disc version 1.3
TITLES
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 20067)
Kitts, P.A.
Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, Clontechn Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT
This sequence has been compiled from information in the sequence
databases, published literature and other sources. If you suspect
there is an error in this sequence, please contact Clontechn's
Technical Service Department at (415) 424-8222 or (800) 662-2566,
extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES
Location/Qualifiers
source
1..20067
/organism="Cloning vector lambda EMBL3"
/mol_type="genomic DNA"
/db_xref="taxon:31783"
ORIGIN
Query Match 82.8%; Score 585.4; DB 12; Length 20067;
Best local similarity 99.8%; Pred. No. 2.3e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 45 ATCTAGCTGATCATGATCATATCGTGGGCTTTTTCGGGCTCATCGCCCAAGC 104

Search Notes

Db 3055 ATCCAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGGCTCATGATGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 164
Db 3115 TGGCGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 3174
QY 165 GCATGGAAGAAGTTTGGCCGAGATGCTGCTGCTGATGAGCGTGAACGAAACGACG 224
Db 3175 GCATGGAAGAAGTTTGGCCGAGATGCTGCTGCTGATGAGCGTGAACGAAACGACG 3224
QY 225 TTATACATGATGATTTGGGAAAGTGTGCGCATGACGCGCTTTAACGTTAACTGTGTT 284
Db 3235 TTATACATGATGATTTGGGAAAGTGTGCGCATGACGCGCTTTAACGTTAACTGTGTT 3294
QY 285 CAGGCCACTGGGATACCAATTCGTCGCGCTTTTCCGGAACAGTTCGGAATGTCGTT 3294
Db 3295 CAGGCCACTGGGATACCAATTCGTCGCGCTTTTCCGGAACAGTTCGGAATGTCGTT 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCGGAACTGCGCGGTGTG 404
Db 3355 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCGGAACTGCGCGGTGTG 404
QY 405 CAGATTATGACAGCGGTGCGCGCTGGATATTAAGTCAAGGAGACGGTATCTGTGC 464
Db 3415 CAGATTATGACAGCGGTGCGCGCTGGATATTAAGTCAAGGAGACGGTATCTGTGC 464
QY 465 TGGATGCCGCAAGATGACATGATATCCCGTGAATTAAGTCAAGGAGACGGTATCTGTGC 3474
Db 3475 TGGATGCCGCAAGATGACATGATATCCCGTGAATTAAGTCAAGGAGACGGTATCTGTGC 3534
QY 525 ATTACGCTTTTGAACCCGCGGAGAGACCGGAGACCTCGCGGTGAATGTGTTTACAGC 584
Db 3535 ATTACGCTTTTGAACCCGCGGAGAGACCGGAGACCTCGCGGTGAATGTGTTTACAGC 584
QY 585 GTGATGAGAGATGAAAGTCTCGACAGCTGCGAAGACAGCAGCT 631
Db 3595 GTGATGAGAGATGAAAGTCTCGACAGCTGCGAAGACAGCAGCT 3641

RESULT 2
XXU02426 20125 bp DNA linear SYN 29-JAN-1997
LOCUS Cloning vector lambda EMBL3 SP6/T7, left arm.
DEFINITION U02426
VERSION U02426.1 GI:413792
KEYWORDS
SOURCE
ORGANISM Cloning vector lambda EMBL3 SP6/T7
REFERENCE Cloning vector lambda EMBL3 SP6/T7
AUTHORS 1 (bases 1 to 20125)
TITLE Kites, P.A.
JOURNAL ClONTECH Vectors On Disc version 1.3
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 20125)
TITLE Kites, P.A.
JOURNAL Direct Submission
COMMENT Submitted (07-OCT-1993) Paul A. Kites, ClONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
This sequence has been compiled from information in the sequence
databases, published literature and other sources. If you suspect
there is an error in this sequence, please contact ClONTECH's
Technical Service Department at (415) 424-8222 or (800) 662-2566,
extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES
Location/Qualifiers
1..20125
/organism="Cloning vector lambda EMBL3 SP6/T7"
/mol_type="genomic DNA"
/db_xref="taxon:31784"

ORIGIN
Query Match 82.8%; Score 585.4; DB 12; Length 20125;
Best Local Similarity 99.8%; Pred. No. 2.3e-148;

Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGGCTCATGATGCCCAAGC 104
Db 3055 ATCCAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGGCTCATGATGCCCAAGC 3114
QY 105 TGGCGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 164
Db 3115 TGGCGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 3174
QY 165 GCATGGAAGAAGTTTGGCCGAGATGCTGCTGCTGATGAGCGTGAACGAAACGACG 224
Db 3175 GCATGGAAGAAGTTTGGCCGAGATGCTGCTGCTGATGAGCGTGAACGAAACGACG 3224
QY 225 TTATACATGATGATTTGGGAAAGTGTGCGCATGACGCGCTTTAACGTTAACTGTGTT 284
Db 3235 TTATACATGATGATTTGGGAAAGTGTGCGCATGACGCGCTTTAACGTTAACTGTGTT 3294
QY 285 CAGGCCACTGGGATACCAATTCGTCGCGCTTTTCCGGAACAGTTCGGAATGTCGTT 3294
Db 3295 CAGGCCACTGGGATACCAATTCGTCGCGCTTTTCCGGAACAGTTCGGAATGTCGTT 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCGGAACTGCGCGGTGTG 404
Db 3355 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCGGAACTGCGCGGTGTG 404
QY 405 CAGATTATGACAGCGGTGCGCGCTGGATATTAAGTCAAGGAGACGGTATCTGTGC 464
Db 3415 CAGATTATGACAGCGGTGCGCGCTGGATATTAAGTCAAGGAGACGGTATCTGTGC 464
QY 465 TGGATGCCGCAAGATGACATGATATCCCGTGAATTAAGTCAAGGAGACGGTATCTGTGC 3474
Db 3475 TGGATGCCGCAAGATGACATGATATCCCGTGAATTAAGTCAAGGAGACGGTATCTGTGC 3534
QY 525 ATTACGCTTTTGAACCCGCGGAGAGACCGGAGACCTCGCGGTGAATGTGTTTACAGC 584
Db 3535 ATTACGCTTTTGAACCCGCGGAGAGACCGGAGACCTCGCGGTGAATGTGTTTACAGC 584
QY 585 GTGATGAGAGATGAAAGTCTCGACAGCTGCGAAGACAGCAGCT 631
Db 3595 GTGATGAGAGATGAAAGTCTCGACAGCTGCGAAGACAGCAGCT 3641

RESULT 3
CVU39284 42529 bp DNA linear SYN 13-APR-1996
LOCUS Cloning vector TlP97-1, lambda phage lacZ translational fusion
DEFINITION U39284
VERSION U39284
KEYWORDS
SOURCE
ORGANISM Cloning vector TlP97-1
REFERENCE Cloning vector TlP97-1, lambda phage lacZ translational fusion
AUTHORS 1 (bases 1 to 42529)
TITLE St Pierre, R. and Linn, T.
JOURNAL A refined vector system for the in vitro construction of
MEDLINE single-copy transcriptional or translational fusions to lacZ
PUBMED Gene 169 (1), 65-68 (1996)
REFERENCE 8635751
AUTHORS 2 (bases 1 to 42529)
TITLE St Pierre, R.
JOURNAL Direct Submission
COMMENT Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
Location/Qualifiers
1..42529
/organism="Cloning vector TlP97-1"
/mol_type="genomic DNA"
/db_xref="taxon:43840"
/note="rmbT2 terminator"

terminator

```
gene      complement(19507. .20760)
          /gene="lacZ"
          complement(19507. .20760)
          /gene="lacZ"
          /codon_start=1
          /product="lactose permease"
          /protein_id="AAC53648.1"
          /db_xref="GI:1066305"
          /translation="MYLKTNTFMWFGLLFFFFPFIFNGAVPPPPPIWLDINHISKSD
          TGIIFPAISLPSLLFOPPLFGLSDKLGLKILMIITGMLWMPAPFFIFIGPLQYN
          ILVGSIVGSIYLCGPNAGAPAVEAEIEKVRNSNFERGARMGCGMAGCAIIVGI
          MFTINOFVFWLGGSCALLAVLFPKATDPSATYANANAGAHSAFSLKALAEIR
          OKRMFTSLYIGVSCITVDVFDQOPANFTSPFATGBGTRFVGVYTMGLLAASIN
          FFAFLINRIGKALLAGTINSVRIIGSSFAISLEAVLILKTLHMEVFPILLVGC
          KYITSOEVRSSATILVCFKPKOLAMIFMSVLAGNMYESTIFGQALVIVGLALGF
          TLISVFLSGPPLSLRQVNEVA"
          complement(20812. .23859)
          /gene="lacZ"
          complement(20812. .23859)
          /gene="lacZ"
          /note="EcoRI site normally present at the 3' end of lacZ
          has been removed by an uncharacterized mutation"
          /codon_start=1
          /transl_table=1
          /product="beta-galactosidase"
          /protein_id="AAC53649.1"
          /db_xref="GI:1066306"
          /translation="VILQKRDMEVNGVTQLNRLAHPFASMRNSEARTPRSEQQR
          SLNGEMRFPAPAEVSEWECDELPAEDTVVVPNNOMHGDAPITNTVYITVN
          PEPVTEPNCVGLTNPVDESMLOEGOTRIIPGVNSAFHLMCGRWVGGQSDRLP
          SEPLSAFLRPGENRLAVMTLRMSDGLSLENDMMRSGTRFVDSILAKPTQSDPH
          VATFNDPFSRAVLBAVONCGELRDYLRVSLMQGTQVASTAPRGGTIDBERG
          YADVTLRLNENRPLWSAEPNLVRAVELATDGLIEAEADVGEREYRINGLL
          LNKCPRLIRGVNHEHPLHQVMDQETQDILLKONFNADVRSHYNHEHMLWT
          LDRGYLVVDEANIETHGVNRLTDDPRMLPMSRVPVSRQNRNHSVILWSL
          GNESGHGNHDAIYRWIKSVDPSPVQVEGGADTTATDIIICPMYAVDEQPPAP
          KMSIKKLSLPGETRPILCEYAHAMNSLGGPKAKYQAFOYRILQGGFPMWDDOS
          LIRKDENGPWSAIGDGDTPNDRQFCMNQLVPADTRPHALTEAHQDQDFQRLS
          GQTEVISEYLFRHSNDELHMVALLDQKPLAAGEVPDVAPOKQKILIELELPQPS
          AGQMLTVRVVQPATWSEAGHISAMQOMLAENLSVTLPAASHAIPHLTSPMDC
          IELGNKQFROSGFLSOWMIGDKKOLLPLRQOFTRAPLDNDIGVSEARIDPNM
          VERKAAGHQAALALQCTADTLADAVLITTAHMOHGTFLFSRKYVINDSGOM
          AITVDVASTDPHPARIGLNCQLAQVAERVNMGLGPOERYPRLTACPDNDL
          SDMTVPVPSBENLRGCTRLNTPGHQWRDIPFNISRSQQLMETSHRLLHABE
          GTWINDGFHMGIGDSDSWSPSABEQLSAGRYHYLWQCK"
          complement(23874. .23957)
          /note="multiple cloning site"
          complement(23973. .24141)
          /note="rmb1 terminator"
          24475. .25665
          /gene="tecr"
          24475. .25665
          /gene="tecr"
          /codon_start=1
          /transl_table=1
          /product="tetraacycline resistance protein"
          /protein_id="AAC53650.1"
          /db_xref="GI:1066307"
          /translation="MKSNNALIVITGVTLDAVIGLVMPVLPGLLRDIVHSDISAH
          YGVLLAYALMOPLCAPVGLISDRGRRLRPRRRLGATIDVIMATPTVYLIVAG
          RIVAGITGATGAVAGATYADITDGEDRRHRLGLMSACGVMGVPAGGLGAIISLH
          APFLAAVUNGVLNLLGCFMLQESHKGERMPLRARPNVPSFPMARGMTVAALMTV
          FFIWQVGVPPALMLVIFGDBRFRMSATMGLSLAVGILHALAQAVTGATRFQF
          KOAIIAGMAADALGVLLAFATRGMAFPIMILLASGGLGMPALQAMISROVDDHGS
          QLOGSLAALSLTSLTGTPLVITAYASASTMGLAMIVGAALVLVCLPLRRAMWR
          ATSR"
          33602. .35450
          /note="imm21 substitution"

mlec_feature
          82.8%; Score 585.4; DB 12; Length 42529;
          Best Local Similarity 99.8%; Pred. No. 2.5e-146;

ORIGIN
Query Match
Best Local Similarity
```

```
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATGTCGGGTCTTTTTCGGGCTCAGTCGCCCAAGC 104
    |||||
DB 3055 ATCCAGCTGCATCAGATCATATGTCGGGTCTTTTTCGGGCTCAGTCGCCCAAGC 3114
QY 105 TGGCGCTATCGGGCATCGGGGAGAAAGAACCCGTGCTTTTCCCGGAGGTTGAAGCG 164
    |||||
DB 3115 TGGGGCTATCGGGCATCGGGGAGAAAGAACCCGTGCTTTTCCCGGAGGTTGAAGCG 3174
QY 165 GCATGAAAGAGTTTGGCGAGATGACTGCTGCTGATTCGATTCGATTCGAGCAAGCAAGC 224
    |||||
DB 3175 GCATGAAAGAGTTTGGCGAGATGACTGCTGCTGATTCGATTCGATTCGAGCAAGCAAGC 3234
QY 225 TTACCATGATGATTCGGGAAAGGTGCGCAATGCAAGCCCTTTAAAGGTGAACCTGTCGT 284
    |||||
DB 3235 TTACCATGATGATTCGGGAAAGGTGCGCAATGCAAGCCCTTTAAAGGTGAACCTGTCGT 3294
QY 285 CAGGCGACCTGGGATACAGATTCGTCGGGCTTTTCCGGAACAGATTCCGGATGTCAGC 344
    |||||
DB 3295 CAGGCGACCTGGGATACAGATTCGTCGGGCTTTTCCGGAACAGATTCCGGATGTCAGC 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCGCGTGTG 404
    |||||
DB 3355 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCGCGTGTG 3414
QY 405 CAGATTATGACAGCGGTGCGGCGCTGGATATTACGTCAGCGAGACGGGTATCTGCGC 464
    |||||
DB 3415 CAGATTATGACAGCGGTGCGGCGCTGGATATTACGTCAGCGAGACGGGTATCTGCGC 3474
QY 465 TGGATGCGCGAAGAAATGACATGATATCCCGTAGATTACCCCGGGGGCGGCTCGTTC 524
    |||||
DB 3475 TGGATGCGCGAAGAAATGACATGATATCCCGTAGATTACCCCGGGGGCGGCTCGTTC 3534
QY 525 ATTACGCTTTTGAACCCGTGAGAGACGGGCGAGACTGCGGTCGAAATGCTTTTACAGC 584
    |||||
DB 3535 ATTACGCTTTTGAACCCGTGAGAGACGGGCGAGACTGCGGTCGAAATGCTTTTACAGC 3594
QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAAGACGCGAGCT 631
    |||||
DB 3595 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAAGACGCGAGCT 3641

RESULT 4
CVU39285 42530 bp DNA linear SYN 13-APR-1996
LOCUS Cloning vector TLp97-2, phage lambda lacZ translational fusion
DEFINITION vector, complete sequence.
ACCESSION U39285
VERSION U39285.1 GI:1066308
KEYWORDS Cloning vector TLp97-2
SOURCE Cloning vector TLp97-2
ORGANISM artificial sequences; vector.
REFERENCE 1 (bases 1 to 42530)
AUTHORS St Pierre,R. and Lim,T.
TITLE A refined vector system for the in vitro construction of
JOURNAL single-copy transcriptional or translational fusions to lacZ
MEDLINE Gene 169 (1), 65-68 (1996)
PUBMED 8635751
REFERENCE 2 (bases 1 to 42530)
AUTHORS Sepiere,R.
TITLE Direct Submision
JOURNAL Submitted (24-OCT-1995) Thomas Lim, Microbiology and Immunology,
UNIVERSITY of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
SOURCE
1. .42530
/organism="Cloning vector TLp97-2"
/mol_type="genomic DNA"
/db_xref="taxon:43841"
complement(19404. .19489)
/note="rmb1 terminator"

terminator
```

```

gene      complement(19507..20760)
          /gene="lacZ"
          /complement(19507..20760)
          /gene="lacY"
          /codon_start=1
          /transl_table=11
          /product="lactose permease"
          /protein_id="AAC53651.1"
          /db_xref="GI:1066310"
          /translation="MYLKNTFMFMFGLFFPFYFIMGAVYPPFPIIMLHINISKSD
          TGIIPASISPLFPLFPLGLSDKLGKLYLMTITGMVMPAPFPIIFIGLLOYN
          LMGSTVGGIYLFPCFNAGAPAVEAFIEKVSRSNFERGARMGCGMALCSIVQI
          MFTINQFVWMLGSCALILAVLFEAKDAPSSATVANAAGNAGSAFLKMLAEFR
          QPLKMLSLVIGVCTYDVPDQOPANFPTSPFATGEGIVFYVTMELNLSIM
          FKAPLIINRIGKNALLAGTISVPIIGSSPATSALEVLTKLHMFVPLVCGF
          KXISQFEVRSATVILVCFCEPKOLAMFMSVAGNMYESIGQGAIVLGLVAGF
          TLISVTLGSPGLSLRQVNEVA"
          /complement(20812..23859)
          /gene="lacZ"
          /complement(20812..23859)
          /gene="lacZ"
          /note="EcoRI site normally present at the 3' end of lacZ
          has been removed by an uncharacterized mutation"
          /codon_start=1
          /transl_table=11
          /product="beta-galactosidase"
          /protein_id="AAC53652.1"
          /db_xref="GI:1066310"
          /translation="VVLQDRDMENPGVTOLNRLAHPAPRSWNESEARTDRPQQLR
          SLNEMRPAMPAPAEVPEBSMLECDLPEDVTVVBSNMQHDAPITYNTVYITVN
          PEPVPTNPQCSYLTENVDSWLOGORRIIFDVGNASFLHMGVWVWGQSDRP
          SEFLDSEFLRAGENELAVVYLHWSGSSYLEDDMMKSGIFPDVSLHKKPTQISDH
          VATFNDPFSRVLAEVOMCGELRDLVATVLSMQGTQVAGTAPGEGIIIDERG
          YADVTLRLINVENPKLMSAFIPVLRVAVELHTAQDIEACDVGFRERYENGL
          LKNGKPLIRGVNRHEHPLHGVNDEQTMVDILMKQNNFNARCSHNPSTIIMSL
          LCGRYGLYVDENAITHGMVPMNRLTDDPMLPMSRVRTVRORRNHPSTIIMSL
          GNSGCHGANDALIRMTKSDPSRPVQEGGADTTATDIIICMYAVADEQEPVAP
          LKIDENGNWSAYGGDFGDTNDPQFQNGIIVADRPHPALTEAHQOQVWVWVS
          GQTIIEVTSYELFRHSNDLHMMVALDCKPLASGEVPLDAPQKQKQILELPEPES
          AGQLMLTVVVOFNATAMSEAGHSAMQMLAENLSVTLPAASHAIPHLTPEMPC
          IELGNKMQFNRQSGFLSOMTIGDKQKQLITLRDQFTRALDNGVSEATRIDSPAM
          VERWKAAGHYQEAALLCTYADTLADAVLITTAHMOHQKTLFI SRKTYRIDSPAM
          AITVDEVASDHPHPARIGLNCOLAQVAFVNMGLGPOENYDPDLTLTACCPMDLPL
          SMDYTPVVRPSENGICRGRLENGPQHMGGDFQFNISRSYQQQLMETSIRHDLHAEE
          GTWLNIDGTHMGIQSDSWSPSVSAEFQLSAGRYHYQLVWQCR"
          /complement(23875..23958)
          /note="multiple cloning site"
          /complement(23974..24142)
          /note="rmbT1 terminator"
          /gene="tetR"
          /gene="tetR"
          /gene="tetR"
          /codon_start=1
          /transl_table=11
          /product="tetracycline resistance protein"
          /protein_id="AAC53653.1"
          /db_xref="GI:1066311"
          /translation="MKSNNALIVLITVTLDAVIGLWMPVLPGILRDIHVSIAH
          YGVLLVLMQFICAPVUGALSDFRGPVYLASLGATIDVIMATTPVIMITVAG
          RIVAGITGATGAGVAGYIADITDGEDRARHGLSACGCVWVAGAGGLAISIAH
          APIAAVANGNANLGLCFMQSHKGRRPMPLRAPNPVSPPMAGRTIVALTAV
          FPIIMQVGVPAALVILFEDEFWRSATMIGLSLVGIGIHALAQAQVTPGATVGE
          KQAITGMAADALGYVILAFATRGWMAPIWILLASGTIMPLQAMLSQVDDHGG
          QLOGSLAALTSLTSLITGPIVTAIVAAASATWNLAVTGAALVYCLPALRGAWSR
          ATST"
          /complement(35451
          /note="1mm21 substitution"
          /complement(35451

```

Query Match 82.8%; Score 585.4; DB 12; Length 42530;
 Best Local Similarity 99.8%; Pred. No. 2.5e-148;

```

Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATGTCGGGCTCTTTTTCGGCTCAGTACGCCCAAC 104
Db 3055 ATCCAGCTGCATCAGATCATATGTCGGGCTCTTTTTCGGCTCAGTACGCCCAAC 104
QY 105 TGGCGCTATCAGGATCAGGATCAGGATCAGGATCAGGATCAGGATCAGGATCAGG 3114
Db 3115 TGGCGCTATCAGGATCAGGATCAGGATCAGGATCAGGATCAGGATCAGGATCAGG 164
QY 165 GCATGGAAGAAGATTTCGCGAGATGACCTGCTGCTGATTTGACGTTGACGAAACGACG 3174
Db 3175 GCATGGAAGAAGATTTCGCGAGATGACCTGCTGCTGATTTGACGTTGACGAAACGACG 3174
QY 225 TTTCACATGATGATTCGGGAGAGGTTGCGCATGACAGCGCTTTAAACGATGATCTGCTT 284
Db 3235 TTTCACATGATGATTCGGGAGAGGTTGCGCATGACAGCGCTTTAAACGATGATCTGCTT 3294
QY 285 CAGGCCACTGGGATACGATTCGTCGCGGCTTTTCCGGAACAGCTTCGGATGTCAGC 344
Db 3295 CAGGCCACTGGGATACGATTCGTCGCGGCTTTTCCGGAACAGCTTCGGATGTCAGC 3354
QY 345 CCGAAGCGATCAGAACCCGGAACATACCGGCGACAGCCGGAACGCTGCGGATGTCAGC 404
Db 3355 CCGAAGCGATCAGAACCCGGAACATACCGGCGACAGCCGGAACGCTGCGGATGTCAGC 3414
QY 405 CAGATTAATGACAGCGGTGCGCGCTGCGGATATTAAGCTGACGAGACGAGTATCTGTC 464
Db 3415 CAGATTAATGACAGCGGTGCGCGCTGCGGATATTAAGCTGACGAGACGAGTATCTGTC 3474
QY 465 TGGATGCGCGAAGATGACATGATACCCCGTGAATTAACCGGCGGCGCGCTGTC 524
Db 3475 TGGATGCGCGAAGATGACATGATACCCCGTGAATTAACCGGCGGCGCGCTGTC 3534
QY 525 ATTACGCTTTTGAACCCGCGTGAAGAGACGCGGAGACCTCGGCGGCGCGCTGTC 584
Db 3535 ATTACGCTTTTGAACCCGCGTGAAGAGACGCGGAGACCTCGGCGGCGCGCTGTC 3594
QY 585 GTGATGAGCAGATGGAAGATGCTCCACACGCTGCGAAGACGACGCT 631
Db 3595 GTGATGAGCAGATGGAAGATGCTCCACACGCTGCGAAGACGACGCT 3641
RESULT 5
LOCUS CUU39286
DEFINITION Cloning vector TUF97-3, DNA linear 42531 bp, SYN 13-APR-1996
ACCESSION U39286
VERSION U39286
KEYWORDS Cloning vector TUF97-3
SOURCE Cloning vector TUF97-3
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 42531)
AUTHORS St Pierre, R. and Linn, T.
TITLE A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lacZ
JOURNAL Gene 169 (1), 65-68 (1996)
MEDLINE 96186904
PUBMED 8635751
REFERENCE 2 (bases 1 to 42531)
AUTHORS Sepiere, R.
TITLE Direct Substitution
JOURNAL Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
source
1..42531
/organism="Cloning vector TUF97-3"
/mol_type="genomic DNA"
/db_xref="taxon:43842"
complement(19404..19489)
/note="rmbT2 terminator"

```

```
gene complement(19507. .20760)
/ gene="lacc"
CDS complement(19507. .20760)
/ gene="lacc"
/codon_start=1
/codon_table=1
/product="laccase permease"
/protein_id="AAC53654.1"
/db_xref="GI:1066313"
/translation="MYLKNITNFMFGLFFFFYPFINGAYPPFPIMLHNDINHISKD
TGLIFPAISLSESLFOPILFGLISDLKLRKLLMIITGLMVPAPFPIFLPGLLOYN
ILVGSIVGIVGLGCPNAGAPAVAEATEKVRNSRNFEGRAMGCGMALCAIVGI
MFTINQFPMWLGSCALLAVLPFAKTDPSSATYANANAGAHNSPSLKALIELR
OPKLMPLSLYVIGSCITYDFDQOPANFTSPATGCGRVPRVYVTMMGLLAASIM
FPAPLITNRIIGKALILAGTIMSVRIIGSSFATSLAEVLVTKLHMPEVFFLLVGP
FYITQFEVRSATIVLCFCFQQLAMIEMSVLAGNMVESIGQAGLVGLVALGP
TLISVFLSGGPLSLRRQYNEVA"
complement(20812. .23859)
/ gene="lacc2"
CDS complement(20812. .23859)
/ gene="lacc2"
/ note="EcoRI site normally present at the 3' end of lacc2
has been removed by an uncharacterized mutation"
/codon_start=1
/codon_table=1
/product="beta-galactosidase"
/protein_id="AAC53655.1"
/db_xref="GI:1066314"
/translation="VVLQRDMENPVYTLNRLAHPFASWRSEBARITDRPSQQLR
SLNEMFAMPAPAEVAVESWLECDLPADTVVVPNMOMHGYDAPYTNVYITTN
PPVPTNPQCYSLTFNVDSEMLQEGOTRIIPGVNSAPHLKMGNGVGGDSRIP
SERPLSFLAGENRLAVMLRMSDGYLEQDMRRSGIFRDVSLHKPTQISDHF
VATFNDPSRAVLEASVOMCGELRDYLRVSLMQEYTOVASITAPFGGIIDERG
YADVTLRLNENPKLMSAEIPNLVAVVELHTADGTLIEACDVGFREVRLENGL
LANKPRLIRGVNHEHPLHQVMEQVLEODILKONNFVRSCHYENHSLMTL
LCDRYGLVVDENAIETHGMVPMNRLTDDPMLPAMSEVLRBMQROHNSVILWL
GNESGHANDALYRMTKSVDSRPNVQEGGADTTATDITCPYARVDEQPPAP
KMSIKKLSLPGETRPLICEYAMKNSLGGFKYQAFQYRILQDGGFVMDVDS
LITKIDENGPASVAGDGFDPNDROFCMGNLVFADTRPHALTEAHQOQFQFRLS
GQITLEVSEYLFRHSNDELHMVNLQDGPLASGEVLDVAPQKQILIELPELQPS
AGOLMLTVRVVQPAATWSEAGHISAMQOMLAENLVLPAAHAIPLHLSMDPC
IELGNKMOFROSGPLSOWMIGDKQLPLRPOFTRAPLNDIGVSEATRIDPMW
VERKAGHQAQEAALIOCTADTLADAVLITTAAMHOGCTPLISKRYRIDSSGM
AITVDEVASDTPHARITGLNQLQVAVERNVIGLQENETPRLTACGRDNDPL
SDMTVPVFPSEBENLRGCTRELNGPOMRQDFQFNISQOQDMETSHRHLHAEB
GTMLNDGFHMGIGDSDSWSPSVAEPQLSAGRHYOLWVCOK"
complement(23876. .23959)
/ note="multiple cloning site"
complement(23975. .24143)
/ note="trnBT1 terminator"
24477. .25667
/ gene="tccr"
24477. .25667
/ gene="tccr"
/codon_start=1
/codon_table=1
/product="tetracycline resistance protein"
/protein_id="AAC53656.1"
/db_xref="GI:1066315"
/translation="MKSNALIVITGTVTLDVAVIGLVMPVLPGLLDIVHSDSIASH
YGVLLAYALMOPLCAPVGLASDRPGRRRPPLASLGATIDVATVATPVYLVING
RIVAGITGATGAVAGATVADITDGEDRRRPHLSSACGCVGAVGAPVAGGLISLH
APFLAAVAVLNGLNLILGCFPLMQESHKGERMPLRAPNPVSPFMARGMTVAILMV
FFIMQVGVPALMLVIFGEBDRFMSATMGLSLAVGGLIHALAQAVTGATRFGE
KQAIILAGMAADALGYLLAFATRCMAFPTMGLIILASGGLIPALQAMLSQVDDHGG
QOGLSLAALTSLSITGPLYTAYASASTMNLAMIVGALVIVCLPLALRRAMSR
ATSR"
misc_feature 33604. .35452
/ note="1mm21 substitution"
```

Query Match 82.8%; Score 585.4; DB 12; Length 42531;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;

```
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATCTCGGGCTTTTTCGGGCTCACTCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGATCATATCTCGGGCTTTTTCGGGCTCACTCGCCCAAGC 3114
QY 105 TGGCGCATCTCGGGCATCGGGGAGGAAGAGCCGTCCTTTTCCCGGAGGTGAAGCG 164
DB 3115 TGGCGCATCTCGGGCATCGGGGAGGAAGAGCCGTCCTTTTCCCGGAGGTGAAGCG 3174
QY 165 GCATGAAGAAGTGTGCGGAGATGACTGCTGCTGATTAAGTGAACGAAACGACG 224
DB 3175 GCATGAAGAAGTGTGCGGAGATGACTGCTGCTGATTAAGTGAACGAAACGACG 3234
QY 225 TTTCATGATGATTCGGGAAGGTGCGCATGACGCTTTTAAAGGTGAACCTGTCGTT 284
DB 3235 TTTCATGATGATTCGGGAAGGTGCGCATGACGCTTTTAAAGGTGAACCTGTCGTT 3294
QY 285 CAGGCCACTGGGATACAGTTCGTCGCGGCTTTTCCGACACAGCTCCGATGTCAGC 344
DB 3295 CAGGCCACTGGGATACAGTTCGTCGCGGCTTTTCCGACACAGCTCCGATGTCAGC 3354
QY 345 CCGAAGCGCATCAGCAACCGAACAATACCGGACGACCGGAACTGCGCGGTGTG 404
DB 3355 CCGAAGCGCATCAGCAACCGAACAATACCGGACGACCGGAACTGCGCGGTGTG 3414
QY 405 CAGATTATGACAGCGGTGCGGCGCTGGATATTACGTCAAGCGAGACGGGTATCTGGC 464
DB 3415 CAGATTATGACAGCGGTGCGGCGCTGGATATTACGTCAAGCGAGACGGGTATCTGGC 3474
QY 465 TGGATGCGCGGAGAAATGGAATGATATCCCGGAGTTACCGCGCGGCGCGCTGTC 524
DB 3475 TGGATGCGCGGAGAAATGGAATGATATCCCGGAGTTACCGCGCGGCGCGCTGTC 3534
QY 525 ATTACGCTTTTGAACCGGTGAGAGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 584
DB 3535 ATTACGCTTTTGAACCGGTGAGAGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 3594
QY 585 GTGATGAGACAGATGAAGATGCTTGACAGCGCTGCAAGACGCGAGCT 631
DB 3595 GTGATGAGACAGATGAAGATGCTTGACAGCGCTGCAAGACGCGAGCT 3641
RESULT 6
LOCUS CVU37692 42704 bp DNA linear SYN 13-APR-1996
DEFINITION Cloning vector lambda TXF97, lacZ transcriptional fusion vector,
complete sequence.
ACCESSION U37692
VERSION U37692.1 GI:1051181
KEYWORDS Cloning vector lambda TXF97
SOURCE Cloning vector lambda TXF97
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 42704)
AUTHORS St Pierre, R. and Linn, T.
TITLE A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lacZ
JOURNAL Gene 169 (1), 65-68 (1996)
MEDLINE 96186904
PUBMED 8635751
REFERENCE 2 (bases 1 to 42704)
AUTHORS St Pierre, R. and Linn, T.
TITLE Direct Submmission
JOURNAL Submitted (29-SEP-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
source
1. 42704
/organism="Cloning vector lambda TXF97"
/mol_type="genomic DNA"
/db_xref="taxon:43300"
1. 42704
/ note="EcoRI site normally present at this site has been
```

```
misc_feature
5505 removed by an uncharacterized mutation"
/Note="BamHI site normally present at this site has been
removed by an uncharacterized mutation"
complement(19404..19489)
/Note="rrnB2 terminator"
gene
complement(19507..20760)
/Note="lacZ"
CDS
complement(19507..20760)
/Note="lacZ"
/codon_start=1
/transl_table=1
/product="lactose permease"
/protein_id="AAC53645.1"
/db_xref="GI:1051182"
/transl_table="MYLKNTVFMFGLEFFFEYFIMGAPPPPIWLDHINISKS
TGIIIPALISLIPOLPLGLSLDKLGRKYLIMTGMLEMPAPPIPTFGULDN
IIVGIVGIVGICFCFNAGAPAVEAIEKVRGRNFEGRAMPCCGMAACASIVGI
MPTINNOVFYMLGSGCALILAVLEFATDAPSSATVAGNANHAFSLKALIFR
OKPLMFLSLVIVGSCYDVDFDOCFANFESFATGCGTRGVYVITMGBELNLSIM
FRAPLINRIGGKNAIILAGTIMSVRIIGSSRAISALBVLILKTLMEFVPLVGC
KITSQFEVRSFATYIVCFEFCFOLMIMSVILAGNMYISIGFQAGYVILGIVLGF
TLISVFTLSGPPSLRLROYNEVA"
complement(20812..23886)
/Note="lacZ"
gene
complement(20812..23886)
/Note="EcoRI site normally present at the end of lacZ has
been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=1
/product="beta-galactosidase"
/protein_id="AAC53646.1"
/db_xref="GI:1051183"
/transl_table="MTMITDSIAVVLQRRDMENPGVTLNRLAHPPEASRWSEAR
TDRPSQQLBSLNGEMRAMPAPAEVAPESWLECDPEADTVVVSNNOMGIDPIYT
NVTYPIVNPBPVPTENPTGCSLTENDESWLQGGOTRII EDGVNAPFLHMCNGVAV
GYGODSRILSEFDSAFELRAGENRLAVULRMSDSYLEDOMMRKSGITRDSALHK
GELIDRGGVADRVTILNENPKLMSAEIPNLYRAVELHLDAGLIEANACDVFR
EVRIENGLLLNGKPLIRGVNHEHHHHCQVADDEOTVODILKONNFANVRSH
YVNHPLMYTLCDRXYGLVVDENAIETTHMPNKLTDPPMLPAMSERVTRMQRORND
HPSVILISGNSGGANHDLIRKISVDPSRVQYEGGADTTATDIIICMYZARND
EDOPFAVKKMSIKKWSLPEIRPLICEVAHAMSGSLGFPKXQVAPROVRIQGG
FYMDWVDOSLIKVDENGNPMSAYGDFDPNDROFOMNLVYADTTPRALTEKHQ
OOEPQRLSGOTIEVTSYLFHSDNLELHMVALQKPLASERVLDIAPRGKOLIE
LBELPESAGQLWTVAVVOPNATAMSEAGHI SAOONRLAENSVTLPAASHAIPH
ATRIDPNAMVERMAAGHYOAEALLOCTADADVILITLAAHQHOKTLIFSRKT
CFRSDMLPLSDMTPPYVPSENGIRGTRRLNQPOMRGDQFPMISRYSOOQMENTS
HRHLHABEGTWINIDGFHMGIGSDSMSVSAEFOLASGRHYQLVWCKK"
complement(24049..24132)
/Note="multiple cloning site"
terminator
complement(24148..24316)
/Note="rrnB1 terminator"
gene
24650..25840
/Note="telR"
24650..25840
/Note="telR"
24650..25840
/Note="Description: tetracycline resistance gene;
tetracycline resistance protein"
/codon_start=1
/transl_table=1
/protein_id="AAC53647.1"
/db_xref="GI:1051184"
/transl_table="MKSNNALIVILGTVTIDAVGIGLVNVPVPLGLRDIIVHSDSIASH
IGVLLALYALQFLCAPVGLSDRGRRRPVLLASLGATIDVAILATTPVIMLIYAG
RIVAGITGATAGVAGVADITDGEDSRRFGLMSCFGGWAGVAGGLIAGATSLH
APFLAAVINGLINLIGCFIMOSHBKSRPMLRPNVSSFRNARSGTIIAALSTH
EFIMQVGOVPAALMVI EGBRRMSATMGLSLAVFGIILALAOAPVTGATKRRGE
KOALIAAGAADALGVYLAIRTRGMAAPFIMGLASGIGMALOAMISROVDDDHQS
OLOGSLAALTSITSITPLIVTAVIYASASTWGLAMVIGALYIVCI PALURKAWSR
```

```
misc_feature
31747 ATST"
/Note="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
misc_feature
33777..35625
/Note="1mm21 substitution"
misc_feature
39168
/Note="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
ORIGIN
Query Match 82.8%; Score 585.4; DB 12; Length 42704;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGATCAGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
DB 3055 ATCCAGCTGCATCAGATCATATCGTCGGGTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
QY 105 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCCGTCCTTTTCCCGAGGTTGAAGC 3114
DB 3115 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCCGTCCTTTTCCCGAGGTTGAAGC 3114
QY 165 GCATGGAAGAGTTTCCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACAG 3174
DB 3175 GCATGGAAGAGTTTCCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACAG 3174
QY 225 TTACCATGATGATTCGGGAAGGTGGCCATGACGCTTTTAAACGATGACGTTGCTT 284
DB 3235 TTACCATGATGATTCGGGAAGGTGGCCATGACGCTTTTAAACGATGACGTTGCTT 3294
QY 285 CAGGCACTCGGATATACAGTTGTCGCGCTTTTCCGAGCAGTTCCGATGCTTCCG 344
DB 3295 CAGGCACTCGGATATACAGTTGTCGCGCTTTTCCGAGCAGTTCCGATGCTTCCG 344
QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACACCGGAACGTCGCGCGGTG 3354
DB 3355 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACACCGGAACGTCGCGGTG 3354
QY 405 CAGATTATACAGCGGTCGCGCTCGGATATTAACGTCAGCGAGACGCGGATACCTGGC 3414
DB 3415 CAGATTATACAGCGGTCGCGCTCGGATATTAACGTCAGCGAGACGCGGATACCTGGC 3414
QY 465 TGGATCCCGCAAAATGACATGATACCCCGTAGTTACCCGCGCGCGCTGCTTC 524
DB 3475 TGGATCCCGCAAAATGACATGATACCCCGTAGTTACCCGCGCGCGCTGCTTC 3474
QY 525 ATTCACGTTTGGAAACCGGTGGAAGAGCGGAGACTTCGCGGTCAATGTGTTTACAGC 584
DB 3535 ATTCACGTTTGGAAACCGGTGGAAGAGCGGAGACTTCGCGGTCAATGTGTTTACAGC 584
QY 585 GGTATGAGCAGATGAAGATGCTCGACACGCTCGACAGAACCGCAGT 631
DB 3595 GTATGAGCAGATGAAGATGCTCGACACGCTCGACAGAACCGCAGT 3641
RESULT 7
LACMG 48502 bp DNA circular PHG 17-APR-2002
LOCUS Bacteriophage lambda, complete genome.
DEFINITION J02459.1 M1723 M24325 V00636 X00906
VERSION J02459.1 GI:215104
KEYWORDS DNA-binding protein; circular; coat protein; complete genome;
SOURCE Bacteriophage lambda
ORGANISM Bacteriophage lambda
Viruses; deDNA viruses; no RNA stage; Caudovirales; Siphoviridae;
Lambda-like viruses.
1 (bases 1 to 12)
REFERENCE Wu, R. and Taylor, E.
AUTHORS Nucleotide sequence analysis of DNA. II. Complete nucleotide
TITLE sequence of the cohesive ends of bacteriophage lambda DNA
JOURNAL J. Mol. Biol. 57 (3), 491-511 (1972)
```

- | | |
|---|---|
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 71209066
4931680
2 (bases 45493 to 45963)
Imada,M. and Tsugita,A.
Amino acid sequence of lambda phage endolysin
Nature New Biol. 233, 230-231 (1971)
3 (sites) |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | Weigel,P.H., Englund,P.T., Murray,K. and Old,R.W.
The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
Proc. Natl. Acad. Sci. U.S.A. 70 (4), 1151-1155 (1973)
73215915
4515613 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 4 (bases 38597 to 38672)
Dahlberg,J.E. and Blattner,F.R.
In vitro transcription products of lambda DNA: Nucleotide sequences
and regulatory sites
(1n) Fox,C.F. and Robinson,W.S. (Eds.)
VIRUS RESEARCH. PROCEEDINGS OF 1973 ICM-UCLA SYMPOSIUM: 533-544;
Academic Press, New York (1973)
5 (bases 37945 to 38027) |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | Maniatis,T., Fasham,M., Backman,K., Kleid,D., Flashman,S.,
Jeffrey,A. and Maurer,R.
Recognition sequences of repressor and polymerase in the operators
of bacteriophage lambda
Cell 5 (2), 109-113 (1975)
75185528
1095210 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 6 (bases 35583 to 35600)
Kleid,D.G., Agarwal,K.L. and Khorana,H.G.
The nucleotide sequence in the promoter region of the gene N in
bacteriophage lambda
J. Biol. Chem. 250 (14), 5574-5582 (1975)
75189495
167018 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 7 (bases 35434 to 35618)
Dahlberg,J.E. and Blattner,F.R.
Sequence of the promoter-operator proximal region of the major
leftward RNA of bacteriophage lambda
Nucleic Acids Res. 2 (9), 1441-1458 (1975)
76031664
1178525 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 8 (bases 37945 to 38018)
Maniatis,T., Jeffrey,A. and Kleid,D.G.
Nucleotide sequence of the rightward operator of phage lambda
Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)
75158212
1055375 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 9 (bases 44588 to 44773)
Sklar,J., Yoc,P. and Weissman,S.M.
Determination of genes, restriction sites, and DNA sequences
surrounding the 6S RNA template of bacteriophage lambda
Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)
75217847
1098044 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 10 (bases 37905 to 37989)
Walz,A., Pirocta,V. and Ineichen,K.
Lambda repressor regulates the switch between PR and P _{RM} promoters
Nature 262 (5570), 665-669 (1976)
76267718
958438 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 11 (bases 37946 to 38039)
Smith,G.R., Eisen,H., Reichardt,L. and Hedgepeth,J.
Deletions of lambda phage locating a prm mutation within the
rightward operator
Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)
76152323
1062780 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 12 (bases 35578 to 35667; 37903 to 38027)
Flashman,M., Backman,K., Humayun,M.Z., Jeffrey,A., Maurer,R.,
Meyer,B. and Sauer,R.T.
Autoregulation and function of a repressor in bacteriophage lambda
Science 194 (4261), 156-161 (1976)
76271154 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 959843
13 (bases 35578 to 35667)
Humayun,Z., Jeffrey,A. and Flashman,M.
Completed DNA sequences and organization of repressor-binding sites
in the operators of phage lambda
J. Mol. Biol. 112 (2), 265-277 (1977)
77209970
875019 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 14 (bases 38610 to 38732)
Scherer,G., Hobom,G. and Kossel,H.
DNA base sequence of the po promoter region of phage lambda
Nature 265 (5590), 117-121 (1977)
77100320
834253 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 15 (bases 38041 to 38241)
Roberts,T.M., Shimatake,H., Brady,C. and Rosenberg,M.
Sequence of Cro gene of bacteriophage lambda
Nature 270 (5634), 274-275 (1977)
78071724
593347 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 16 (bases 27616 to 28935)
Davies,R.W., Schreier,P.H. and Buchel,D.E.
Nucleotide sequence of the attachment site of coliphage lambda
Nature 270 (5639), 757-760 (1977)
78071823
593399 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 17 (bases 37206 to 37263; 37914 to 37970)
Humayun,Z.
DNA sequence at the end of the CI gene in bacteriophage lambda
Nucleic Acids Res. 4 (7), 2137-2143 (1977)
78011659
909767 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 18 (bases 27617 to 27934)
Landy,A. and Rose,W.
Viral integration and excision: structure of the lambda att sites
Science 197 (4309), 1147-1160 (1977)
77258934
331474 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 19 (bases 38062 to 39170)
Denniston-Thompson,K., Moore,D.D., Kruger,K.E., Furch,M.E. and
Blattner,F.R.
Physical structure of the replication origin of bacteriophage
lambda
Science 198 (4321), 1051-1056 (1977)
78054731
929187 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 20 (bases 44467 to 44807)
Sklar,J.L.
Structure and function of two regions of DNA controlling the
synthesis of prokaryotic RNAs
Thesis (1977)
21 (sites) |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | Adhya,S. and Gottesman,M.
Control of transcription termination
Annu. Rev. Biochem. 47, 967-996 (1978)
78234064
354508 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 22 (bases 13 to 72; 48391 to 48502)
Nichols,B.P. and Donelson,J.E.
178-Nucleotide sequence surrounding the cos site of bacteriophage
lambda DNA
J. Virol. 26 (2), 429-434 (1978)
78197067
666898 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 23 (bases 37938 to 38016; 35589 to 35666)
Flashman,S.M.
Mutational analysis of the operators of bacteriophage lambda
Mol. Gen. Genet. 166 (1), 61-73 (1978)
79114073
368570 |
| JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE | 24 (bases 37990 to 38982)
Schwarz,B., Scherer,G., Hobom,G. and Kossel,H.
Nucleotide sequence of cro, cII and part of the O gene in phage |

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	1 Lambda DNA Nature 272 (5652), 410-414 (1978) 78135462 264238 25 (bases 38212 to 38362) Rosenberry,M., Court,D., Shmatake,H., Brady,C. and Wulff,D.L. The relationship between function and DNA sequence in an interictronic regulatory region in phage lambda Nature 272 (5652), 414-423 (1978) 78135463 634366 26 (bases 37224 to 37940) Sauer,R.T. DNA sequence of the bacteriophage gamma ci gene Nature 276 (5685), 301-302 (1978) 79053284 714163 27 (bases 38597 to 39688) Scherer,G. Nucleotide sequence of the O gene and of the origin of replication in bacteriophage lambda DNA Nucleic Acids Res. 5 (9), 3141-3156 (1978) 704348 28 (bases 29711 to 29811: 31043 to 31058) Davies,R.W., Schreiber,P.H. and Buchel,D.E. Determination of the endpoints of partial deletion mutants of the attachment site of bacteriophage lambda by DNA sequencing Nucleic Acids Res. 5 (9), 3209-3218 (1978) 79033246 704352 29 (bases 21661 to 31129) Hoess,R.H. and Landy,A. Structure of the lambda att sites generated by int-dependent deletions Proc. Natl. Acad. Sci. U.S.A. 75 (11), 5437-5441 (1978) 79074833 364480 30 (bases 38453 to 38500)
---	--

Db 3475 TGGATCCCGCAGAAATGCAATGATATCCCGCGAGTTTACCCGGCGGCGCTCGTTC 353

Qy 525 ATTCAGCTTTTGAACCCGTGGAGACGGCAGACTCGGTGCAAAATGTTTACAG 584

Db 3535 ATTCAGCTTTTGAACCCGTGGAGACGGCAGACTCGGTGCAAAATGTTTACAG 359

Qy 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGCT 631

Db 3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGCT 3641

RESULTS	9	
AF434923		
LOCUS		
DEFINITION	5969 bp	DNA
ACCESSION	AF434923	circular
KEYWORDS	AF434923.1	GI.16904152
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		

Query Match	82.8%	Score 585.4	DB 7	Length 48502
Best Local Similarity	99.8%	Pred. No. 2,5e-148		
Matches 586	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	45	ATTCAGTCGATCAGATTCATATGATGGGCTCTTTTTCCTGCTCAGTCATGCCCCAAGC	104	
Db	3055	ATCCAGCTGCATCAGATTCATATGTCGGCTCTTTTTCCTGCTCAGTCATGCCCCAAGC	3114	
QY	105	TGGCGCTATCGGGCATCGGGGAGAAAGGCCGTGCTTTTTCCTGCGAGGTTGAAGCG	164	
Db	3115	TGGCGCTATCTGGGCATCGGGGAGAAAGGCCGTGCTTTTTCCTGCGAGGTTGAAGCG	3174	
QY	165	GCATGGAAGAAGTTTGC CGAGATGCACTGCTCTCATTTGACGTTGAGCCGAAAACGACG	224	
Db	3175	GCATGGAAGAAGTTTGC CGAGATGCACTGCTCTCATTTGACGTTGAGCCGAAAACGACG	3234	
QY	225	TTTAAACATGATGATTTGCGAAGGTGGCATGCAACGCTTTAAACGCTGAACTGTTGGTT	284	
Db	3235	TTTAAACATGATATTCGGGAAGGTGGCATGCAACGCTTTAAACGCTGAACTGTTGGTT	3294	
QY	285	CAGGCCACCTGGGAATACCAAGTTTCGTGCGGCTTTTTCGGAACAAGTTCCGGAATGGTCACG	344	
Db	3295	CAGGCCACCTGGGAATACCAAGTTTCGTGCGGCTTTTTCGGAACAAGTTCCGGAATGGTCACG	3354	
QY	345	CCGAAGCCGATACGAACCCGAACAATACCGGACGACGCGGAACCTGCGTCGGGTGTG	404	
Db	3355	CCGAAGCCGATACGAACCCGAACAATACCGGACGACGCGGAACCTGCGTCGGGTGTG	3414	
QY	405	CAGATTAAATGACAGCGGTGGCGGCTCTGGGATATTACGTACGACGAGGACGGGTATCTGGC	464	
Db	3415	CAGATTAAATGACAGCGGTGGCGGCTCTGGGATATTACGTACGACGAGGACGGGTATCTGGC	3474	
QY	465	TGATGCGCGCAAAATGACATGAAATACCCCGTAGATTACCGGGGGGCGCGCTCGTTC	524	

enhancer	Promoter	CDS	misc_feature
<p>97. 614</p> <p>/note="Autographa californica nuclear polyhedrosis virus h5 enhancer"</p> <p>615. 1201</p>	<p>1328. 2980</p> <p>/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"</p> <p>1328. 2980</p> <p>/codon_start=1</p> <p>/product="firefly luciferase"</p> <p>/protein_id="AA130778.1"</p> <p>/db_xref="GI:16904153"</p> <p>/translation="MEDARKINKKGPAPPPLEDDGAGBQIHKRANKYALVPGTIAFTD AHIENTVTAAYEWMVRALBAMKRYGANTNHRIVVCSNSLDFEMFVGLAFIGAAV APANDIVNERELLSNNSIOPVTVFVSSKGGIQLINQKLPITIOKIIMDSKTDYOG FSHARDIPFNQOIIIPDTAITSVPFHGFGFTLLGYLICGFVLMVFEELSLRS LODYKIOALIVPLTPEFPASTSLIDKDYLSNLSHEIASGAPLSEKVGAAVRRPLD GIRGCVGLTETTSATILPEBGDKRGAAKVVPFEAKVVDLDPTGKTGSVNRGSLP RGPIMWGVYNNFEATNALIDKDGWLSGSIAYWDEDEHFTVIDRLKSLIKTKGQVA PALEHSLIQHNPIEDPAGAGLPPDDAGSELPAVVVLEHGKMTPEKELVIDVASVTT AKTLRGGVAVDEVRKGLTGLDLRKIREILIRAKKGKSKL"</p> <p>3146. 3351</p>	<p>1138. 2980</p> <p>/note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"</p> <p>4168. 5028</p> <p>/codon_start=1</p> <p>/product="beta-lactamase"</p> <p>/protein_id="AA130779.1"</p> <p>/db_xref="GI:16904154"</p> <p>/translation="MSIQHRRVALIPFPAFCPLPEAHPETLNVKXDAEQUGARVGY FILDNSKLTIESFRPEERPMMSSTFKVLCAVSIIRIDACQDGLGRIRHNSQNDIVE YSPVTERKLTGDMTVRELCSAAITMSDNTAANULLTTTGGKELTALAHNMGGNDVRL</p>	<p>1138. 2980</p> <p>/note="Heliothis virescens p63 chaperonin gene double poly-adenylation sequence"</p> <p>4168. 5028</p> <p>/codon_start=1</p> <p>/product="beta-lactamase"</p> <p>/protein_id="AA130779.1"</p> <p>/db_xref="GI:16904154"</p> <p>/translation="MSIQHRRVALIPFPAFCPLPEAHPETLNVKXDAEQUGARVGY FILDNSKLTIESFRPEERPMMSSTFKVLCAVSIIRIDACQDGLGRIRHNSQNDIVE YSPVTERKLTGDMTVRELCSAAITMSDNTAANULLTTTGGKELTALAHNMGGNDVRL</p>

DRPEPELNEAIPNDERDTPVAMATTLRLKLTJGBLLTLASROOLIDMMEADKXVAGPL
 LRSALPMGFIAADSGAGERSRGIIALGSDGPRSRVIVYTTGSOATMDERKQIA
 EIGASLTKHM"
 5316..5735
 rep_origin /note="ColE1"

Query Match 81.3%; Score 574.6; DB 12; Length 5969;
 Best Local Similarity 99.3%; Pred. No. 1.7e-145;
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAAAGC 104
 DB 3508 ATCCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAAAGC 3567
 QY 105 TGGCGCTATCTGGGCTATCGGGAGAGAACCCCGTGGCTTTTTCGGGAGTTGAACG 164
 DB 3568 TGGCGCTATCTGGGCTATCGGGAGAGAACCCCGTGGCTTTTTCGGGAGTTGAACG 3627
 QY 165 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 224
 DB 3628 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 3687
 QY 225 TTTCACATGATGATTCGGGAGAGTGTGGCCATGACGCGCTTTAAACGGTGAAGTTCGTT 284
 DB 3688 TTTCACATGATGATTCGGGAGAGTGTGGCCATGACGCGCTTTAAACGGTGAAGTTCGTT 3747
 QY 285 CAGGCCACTGGGATACCAAGTTTCGCGGGCTTTTTCGGAGACAGTTCCGAGTGTGACG 344
 DB 3748 CAGGCCACTGGGATACCAAGTTTCGCGGGCTTTTTCGGAGACAGTTCCGAGTGTGACG 3807
 QY 345 CCGAAGGCGATACCAACCCGAAACATACCGGCGACGACCGGAACTGCCGTGCGGCTGTG 404
 DB 3808 CCGAAGGCGATACCAACCCGAAACATACCGGCGACGACCGGAACTGCCGTGCGGCTGTG 3867
 QY 405 CAGATTAATGACAGCGGTGCGGCTGAGATATTACGTCACGAGAGACGGGTATCTTGAC 464
 DB 3868 CAGATTAATGACAGCGGTGCGGCTGAGATATTACGTCACGAGAGACGGGTATCTTGAC 3927
 QY 465 TGGATGCCGAGAAATGACATGATGATACCCGCTGATGATACCCGCGGCGGCGCTCGTT 524
 DB 3928 TGGATGCCGAGAAATGACATGATGATACCCGCTGATGATACCCGCGGCGGCGCTCGTT 3987
 QY 525 ATTACAGTTTGTGAACCCGTGAGAGACGGGACGACTCGCGGTGCAATGTGTTTACACG 584
 DB 3988 ATTACAGTTTGTGAACCCGTGAGAGACGGGACGACTCGCGGTGCAATGTGTTTACACG 4047
 QY 585 GTGATGAGACGATGATGATGCTGACACGCTGACGAAAC 625
 DB 4048 GTGATGAGACGATGATGATGCTGACACGCTGACGAAAC 4088

RESULT 9
 AF434925 6182 bp DNA circular SYN 04-APR-2002
 LOCUS AF434925
 DEFINITION Expression vector 409-FOR, complete sequence.
 ACCSSION AF434925
 VERSION AF434925.1 GI:16904158
 KEYWORDS
 SOURCE
 ORGANISM
 Expression vector 409-FOR
 Expression vector 409-FOR
 artificial sequences; vectors.
 REFERENCE
 1 (bases 1 to 6182)
 Zieles, H. and Huynh, C.O.
 Inton-dependent stimulation of marker gene expression in cultured
 insect cells
 Insect Mol. Biol. 11 (1), 87-95 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 21830665
 11841506
 2 (bases 1 to 6182)
 Zieles, H. and Huynh, C.O.
 Direct Submission
 JOURNAL
 Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of

Parasitic Diseases, National Institutes of Health, 4 Center Drive
 MSC 0425, Bethesda, MD 20892-0425, USA
 Location/Qualifiers

FEATURES

source
 1..6182
 /organism="Expression vector 409-FOR"
 /mol_type="genomic DNA"
 /db_xref="taxon:176967"
 /note="insect-specific"
 97..614
 /note="Autographa californica nuclear polyhedrosis virus
 h5 enhancer"
 615..1201
 /note="Autographa californica nuclear polyhedrosis virus
 immediate early promoter"
 1266..1398
 /note="artificial intron"
 1541..3193
 /codon_start=1
 /product="firefly luciferase"
 /protein_id="AAJ30782.1"
 /db_xref="GI:16904159"
 /translation="MEDAKNIKKGPAPFPLDGTAGEOLHKAMRYALVPGTIAFTD
 AHLENTTVAEYFEMSVRLAEAMRYGLNTNHRIVVSESLQFEMFVLGALFTGVAV
 APANDIVNERELNSNMI SQTPTVFSKGLQKLIYOKLPILOKLIIMDSKTYOG
 POSWYTFVTHSLPRGFNDVFPVESPDRDITLIMSSSTGI PKVALPHRTACVR
 FSHARDPITFNGIITPDAIISVPPHHGPFMTTLGLIGFPRVLYMRPEELPLS
 LQDYKIOSALVPLTFSPFAKSLIDRYDLSNLEISGAPLSKEGEAAAKFPLP
 GIRGYVLTSTTSAILITPBGDDKPGAVGVNPFPEKAVVDLDGKLTLYGVQRELCV
 RGPIMSGVYNNPEATNALIDKDMISGDIAYVDEDEHFFIVRLSLIKYQVQA
 PAELIESILPHNPIDAGVAGLPDDDGELPAVAVVLEHGKTMKEKIVDVASQVTT
 AKLRGVVFPVDEVPKGLTKLDAKRLRLIYAKKGRKSL"
 3359..3464
 /note="Heliothis virescens p63 chaperonin gene double
 poly-adenylation sequence"
 4381..5241
 /product="beta-lactamase"
 /protein_id="AAJ30783.1"
 /db_xref="GI:16904160"
 /translation="MSIOHPRVALIIPFPAFCULPVAPHETLVKVKDAEDQLGARVY
 IEIDLNGSKILSFPRPEPRPMSFTVLLCGAVLSRIDGQELGRIRHYSQNDLVE
 YSPVTEGHTLDGQVRELCSAALTMSNTNANLLITLGPKELTFLHNMGDHVTPL
 DRPEPELNEAIPNDERDTPVAMATTLRLKLTJGBLLTLASROOLIDMMEADKXVAGPL
 LRSALPMGFIAADSGAGERSRGIIALGSDGPRSRVIVYTTGSOATMDERKQIA
 EIGASLTKHM"
 5529..5948
 /note="ColE1"

ORIGIN

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
 Best Local Similarity 99.3%; Pred. No. 1.7e-145;
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAAAGC 104
 DB 3721 ATCCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAAAGC 3780
 QY 105 TGGCGCTATCTGGGCTATCGGGAGAGAACCCCGTGGCTTTTTCGGGAGTTGAACG 164
 DB 3781 TGGCGCTATCTGGGCTATCGGGAGAGAACCCCGTGGCTTTTTCGGGAGTTGAACG 3840
 QY 165 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 224
 DB 3841 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 3900
 QY 225 TTTCACATGATGATTCGGGAGAGTGTGGCCATGACGCGCTTTAAACGGTGAAGTTCGTT 284
 DB 3901 TTTCACATGATGATTCGGGAGAGTGTGGCCATGACGCGCTTTAAACGGTGAAGTTCGTT 3960
 QY 285 CAGGCCACTGGGATACCAAGTTTCGCGGGCTTTTTCGGAGACAGTTCCGAGTGTGACG 344
 DB 3961 CAGGCCACTGGGATACCAAGTTTCGCGGGCTTTTTCGGAGACAGTTCCGAGTGTGACG 4020

Qy 345 CCGAAGCCGATCAGCAACCCGAAATATCCGCGACAGCCGGAATGCGCGTCCGGTGG 404
Db 4021 CCGAAGCCGATCAGCAACCCGAAATATCCGCGACAGCCGGAATGCGCGTCCGGTGG 4080
Qy 405 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGAGACGGGTATCTGGC 464
Db 4081 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGAGACGGGTATCTGGC 4140
Qy 465 TGGATGCCGAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGGCGCGCTCGTTC 524
Db 4141 TGGATGCCGAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGGCGCGCTCGTTC 4200
Qy 525 ATTACAGTTTATGACCCCGTGAAGAGACGGGACAGCTCCGCGGTGAGAAATGTTTATACAGC 584
Db 4201 ATTACAGTTTATGACCCCGTGAAGAGACGGGACAGCTCCGCGGTGAGAAATGTTTATACAGC 584
Qy 585 GTGATGAGACAGATGAAGATGCTGACACGCTGCGAAGAC 625
Db 4261 GTGATGAGACAGATGAAGATGCTGACACGCTGATGATAAC 4301

RESULT 10
AF434926 6182 bp DNA circular SYN 04-Apr-2002
LOCUS AF434926
DEFINITION Expression vector 409-MUT, complete sequence.
ACCESSION AF434926
VERSION AF434926.1 GI:16904161
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1. 6182
/organism="Expression vector 409-MUT"
/mol_type="genomic DNA"
/db_xref="taxon:176968"
/note="insect-specific"
97.614
/note="Autographa californica nuclear polyhedrosis virus
h75 enhancer"
615..1201
/note="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
1266..1398
/note="artificial intron with mutated 3' splice site"
1541..3193
/codon_start=1
/product="firefly luciferase"
/protein_id="AAJ30784.1"
/translation="MEDAKNIKKGPAPYPLEDTAGBOLHKKMKRYALVPGTIAFTD
AIPNDIYERELNSMNIOPVVFVSKGOKILNVOKPLIIOKIIIMDSKIDYOG
FQSWYTFVSHLPFGFNEDYDVPSPRDKTIALINSSGSLPGVVALPRTACVR
LQDYKIQSALLVPLTFSPRSTLIDYVFNHGFQFTLGLICGFVVALPFEELPFRS
GIRQYGLTETTSALITPBGDDPRGAVGVVFPFAKVVVDLDTGKLGAVNRPLP
RGPMTGMYVNNPFTALIDKGMHSGDIAVVDDEHPITVDTKSLIKIKGVOA
PAIESIILQHPNIFDAGVAGLPDDDAGELPAVVDLDEHGTETREIVDVASQVTT
AKKLRGVVFPDEVKGLTGKIDARKIRBELILIAKKKGSKL"

promoter
intron
CDS

misc_feature 3359..3464
/note="Heliothis virescens p63 chaparonin gene double
poly-adenylation sequence"
4381..5241
CDS
/codon_start=1
/product="beta-lactamase"
/protein_id="AAJ30785.1"
/db_xref="GI:16904163"
/translation="MSIQHFRVALIPEPAFLPVFAHPETLVKVKDAEDQLGARVGY
YSFVTEKHLDGMTVAELCSAAITMSDNTANILITIGSRKELFLHMGDVTLE
LRSAIPAGWFIADKSGAGRGSRGIIAALGPDGRSRIVLYITGSAITDERRRQIA
EIGASLIKHW"
rep_origin
/note="ColEI"

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.7e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGCTCATGTCATGCCCAAC 104
Db 3721 ATCTAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGCTCATGTCATGCCCAAC 3780
Qy 105 TGGCGCTATCTGGGCTATGCGGAGAGAACCCGCTGCTTTTCCGCGAGTTGAAGC 164
Db 3781 TGGCGCTATCTGGGCTATGCGGAGAGAACCCGCTGCTTTTCCGCGAGTTGAAGC 3840
Qy 165 GCATGAAAGATTGGCCGAGATGACTGCTGCTGATGAGTGAAGCAAAACGACAG 224
Db 3841 GCATGAAAGATTGGCCGAGATGACTGCTGCTGATGAGTGAAGCAAAACGACAG 3900
Qy 225 TTACCATGATGATGCGGAGAGGTGGCCATGACAGCCTTTAAGCGTGAAGTGTGTT 284
Db 3901 TTACCATGATGATGCGGAGAGGTGGCCATGACAGCCTTTAAGCGTGAAGTGTGTT 3960
Qy 285 CAGGCCACCTGGGATACAGTTGCTGCGGCTTTTCCGACACAGTTCGGATGTCAGC 344
Db 3961 CAGGCCACCTGGGATACAGTTGCTGCGGCTTTTCCGACACAGTTCGGATGTCAGC 4020
Qy 345 CCGAAGCCGATCAGCAACCCGAAATATCCGCGACAGCCGGAATGCGCGTCCGGTGG 404
Db 4021 CCGAAGCCGATCAGCAACCCGAAATATCCGCGACAGCCGGAATGCGCGTCCGGTGG 4080
Qy 405 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGACGGGTATCTGGC 464
Db 4081 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGACGGGTATCTGGC 4140
Qy 465 TGGATGCCGAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGGCGCGCTCGTTC 524
Db 4141 TGGATGCCGAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGGCGCGCTCGTTC 4200
Qy 525 ATTACAGTTTATGACCCCGTGAAGAGACGGGACAGCTCCGCGGTGAGAAATGTTTATACAGC 584
Db 4201 ATTACAGTTTATGACCCCGTGAAGAGACGGGACAGCTCCGCGGTGAGAAATGTTTATACAGC 584
Qy 585 GTGATGAGACAGATGAAGATGCTGACACGCTGCGAAGAC 625
Db 4261 GTGATGAGACAGATGAAGATGCTGACACGCTGATGATAAC 4301

RESULT 11
AF434927 6182 bp DNA circular SYN 04-Apr-2002
LOCUS AF434927
DEFINITION Expression vector 410-FOR, complete sequence.
ACCESSION AF434927
VERSION AF434927.1 GI:16904164
KEYWORDS
SOURCE
ORGANISM
Expression vector 410-FOR
Expression vector 410-FOR
artificial sequences; vectors.

REFERENCE 1 (bases 1 to 6182)
AUTHORS Ziebler, H. and Huynh, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Ziebler, H. and Huynh, C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
Source
1..6182
/organism="Expression vector 410-FOR"
/mol_type="genomic DNA"
/db_xref="taxon:176969"
/note="insect-specific"
enhancer
97..614
/note="Autographa californica nuclear polyhedrosis virus h75 enhancer"
promoter
615..1201
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
CDS
1328..12980
/codon_start=1
/product="firefly luciferase"
/protein_id="AAJ30786.1"
/db_xref="GI:16904165"
/translation="MEDAKNTKKGPAPYPLEDNGRQLHKMKRYALVPTIAPTD AHIEVNTIYAYIFEMSVRLAEMKRYGANTHRIYVCEENSLQFFMPYLGLPIGVAV APANDIYNERELNSWNTISQPTVAVSKKQKILNVQKPIIOKITIMDSKTDOG FQSHADPFTVSHLPGEFNEYDFVPSFDDKTIALIMSSGSGLPKVALPHRTACVR PSHARDPIFGNOIIPDTAILSVFPFHGFGFTTIGYLICGRVILVFESEELFELS LODYKIGALVPLTFSFPAKSTLIDKDLNLHIEAGGAPLSKEVEGAAKXPHLP GIDRGYGLTETTSAILITPEBGDDKRGANVPPRAVVDLDGKTXGIVNORIELCY RGRPMISGYVNNPREATNALIKDCKLHSGDIAYNDEDFPTVDRLKSLIKRYQVA PALESLILQHPIFDAGVAGLPDDDAEGLPAAVVLEHGTMEKETVDVAVSOVTT AKKURGVAVFVDEVPKGLTGKLDARKIREILIKAKGKSKL"
intron
3196..3328
/note="artificial intron"
3359..3464
/note="heliothis virescens p63 chaperonin gene double poly-adenylation sequence"
CDS
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAJ30787.1"
/db_xref="GI:16904166"
/translation="WSIQHFRVALIPFPAAFLCLPVFAHPETLVKVKADQDGLARVGY IELDLSGKILIESFRPERFPFMSTFPAKVLGAVLSRIDAOEQDQGRIRHNSQNDLVE YSPVTEKHLTDGKTVRELCSAATMSDNTANLLLTIGKPEKELTAFIHNMGDHYTRL DRWEPELNEAIPNDERDPTMVAATTKRLTGSELTLIASRQOLIDMEADKVAGPL LRSALPGWFIADKSGAGERSRGIALLGPDGKRSRLVIVYTTGSQATMDERRRQILA EIGASLIKHW"
FEATURES
1..5948
/note="ColE1"
rep_origin

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.7e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

45 ATCTAGTCATCGATCATATCGTGGGCTTTTTCGAGCTCATGATCGCCCAAGC 104
DB ATTCAGGCGATCGATCATATCGTGGGCTTTTTCGAGCTCATGATCGCCCAAGC 3780
105 TGGGCGCTATCTGGGCGATCGGAGAGAGAGCCGCTCTTTCGCGAGAGTTGAACG 164
DB TGGGCGCTATCTGGGCGATCGGAGAGAGAGCCGCTCTTTCGCGAGAGTTGAACG 3840
165 GCATGGAAGAGTTTCCGAGAGATGACCTGCTGCATTGACGTTGAGCAAAACGACG 224

DB 3841 GCATGGAAGAGTTTCCGAGAGATGACCTGCTGCATTGACGTTGAGCAAAACGACG 3900
QY 225 TTTCACATGATGATTCCGGAAGAGTGTGGCCATGACCGCTTTAAACGTGTTCTGTT 284
DB 3901 TTTCACATGATGATTCCGGAAGAGTGTGGCCATGACCGCTTTAAACGTGTTCTGTT 3960
QY 285 CAGGCCACCTGGGATACCAATTGCTGGGCGCTTTTCCGGACACAGTTCCGGATGTCACG 344
DB 3961 CAGGCCACCTGGGATACCAATTGCTGGGCGCTTTTCCGGACACAGTTCCGGATGTCACG 4020
QY 345 CCGAAGCGCATCAGAACCCGGAACAATACCGGCGACGCGGAACTGCGCGGTGTG 404
DB 4021 CCGAAGCGCATCAGAACCCGGAACAATACCGGCGACGCGGAACTGCGCGGTGTG 4080
QY 405 CAGATTAAATGACACGCGGTGGCGCGCTGGATATTACGTACGACGAGACGAGTATCTGGC 464
DB 4081 CAGATTAAATGACACGCGGTGGCGCGCTGGATATTACGTACGACGAGACGAGTATCTGGC 4140
QY 465 TGGATGCCGCAAAATGACATGATACCCCTGATGATTACCCGCGGCGCGCTCTGTTG 524
DB 4141 TGGATGCCGCAAAATGACATGATACCCCTGATGATTACCCGCGGCGCGCTCTGTTG 4200
QY 525 ATTCACGTTTGTGAACCGGTGAGAGACGCGGACGCTGGGATGAATGTGTTTACAGC 584
DB 4201 ATTCACGTTTGTGAACCGGTGAGAGACGCGGACGCTGGGATGAATGTGTTTACAGC 4260
QY 585 GTGATGAGACAGATGAAGATGCTGACACGCTGAGAAAC 625
DB 4261 GTGATGAGACAGATGAAGATGCTGACACGCTGAGAAAC 4301
RESULT 12
LOCUS AF434931 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 409-REV, complete sequence.
ACCESSION AF434931
VERSION AF434931.1 GI:16904176
KEYWORDS
SOURCE
Expression vector 409-REV
Expression vector 409-REV
artificial sequences; vectors.
REFERENCE
1 (bases 1 to 6182)
AUTHORS Ziebler, H. and Huynh, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Ziebler, H. and Huynh, C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
Source
1..6182
/organism="Expression vector 409-REV"
/mol_type="genomic DNA"
/db_xref="taxon:176973"
/note="insect-specific"
enhancer
97..614
/note="Autographa californica nuclear polyhedrosis virus h75 enhancer"
promoter
615..1201
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
intron
1541..3193
/codon_start=1
/product="firefly luciferase"
/protein_id="AAJ30794.1"

```

/db_xref="GI:16904178"
/translation="MEDAKNIKGPAPFPYPLEDGTAGEOLHKAMKRYALVPGTIAFTD
AHENVTITAEYFEWVSRLAEAMKRYGLANTNHRIVUSSENSLOFEMFVGALEIVAV
APANDIYNRELNSMNIISOPTVVFVSKGOKLIVNOKPLTIQKIIIMDSKTVOG
FOSWTFVTSHLPGFNEVDFPESPRDKTIALIMSSGSGTLPGVALPRTACVR
FSHARDPIFGNOIIPDTAILSVPHHGFEMFTLGYLICGFVVLMTAFEBELPLRS
LQDYKIQSALVPTLFSFPASTIIDLKYLDSNHEIASGAPLSKEVGAVKRPHLP
GIRQYGLTETTSAILITPESGDRPGAVGVVPEFAKVVDLDTGKTIGVNRGELCV
RGPIMSGYVNNPRATNALIDKQWMLSGDIAVWDEDEHFFIVDLKSLIKKGYOVA
PAELSTILLOPHNIFDAGVAGLPDDAGELPAVVVLEHGKTMKEKELVIVVASQVTT
AKKIRGAVVFVDEVPKGLTGKLDARKIREILIKAKKGSKL"
3529..5948
misc_feature
/feature="Heliothis virescens p63 chaparonin gene double
poly-adenylation sequence"
4381..5241
CDS
/codon_start=1
/product="beta-lactamase"
/protein_id="AAJ30795.1"
/db_xref="GI:16904178"
/translation="MSIQHFRVALIPPPAFCPLVPNHPETLVYVKDAEDQLGARVGY
YSPVTEKHLTDGMYRELCSAIIWTSNDTANILITLIGPKELTAFILNMGDHYRL
DRWEPELNEAIPIIDREDITMPVMAATTLRIKLITGELLTLASRQQLIDWMEADKVAEPL
EIGASLIKHM"
5529..5948
rep_origin
/feature="ColE1"

```

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
 Best Local Similarity 99.3%; Pred. No. 1.7e-145;
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

45 ATCTGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTATGCCCAAGC 104
3721 ATTCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTATGCCCAAGC 104
105 TGGGCTATCTGGGCTGCAGGAGAGAAAGCCCGTCTTTCCGGGAGATTGAAGCG 164
3781 TGGGCTATCTGGGCTGCAGGAGAGAAAGCCCGTCTTTCCGGGAGATTGAAGCG 164
165 GCATGGAAGAAGTTTCCGAGATGACTGCTGCTGCTTAAAGCTTAAAGCGAAACG 224
3841 GCATGGAAGAAGTTTCCGAGATGACTGCTGCTGCTTAAAGCTTAAAGCGAAACG 224
225 TTTCACATGATGATTCGGGAAGGTGTGGCATGCAAGCTTTAAAGCTTAAAGCTT 284
3901 TTTCACATGATGATTCGGGAAGGTGTGGCATGCAAGCTTTAAAGCTTAAAGCTT 284
285 CAGGCCACTGGGATACCAAGTTCTGCGGCTTTTCCGGGACACAGTTCCGATGCT 3960
3961 CAGGCCACTGGGATACCAAGTTCTGCGGCTTTTCCGGGACACAGTTCCGATGCT 3960
345 CCGAAGCCCATCAGCAACCGGAACAATCCGGGACAGCCGGAACCTGCGGATGCT 4020
4021 CCGAAGCCCATCAGCAACCGGAACAATCCGGGACAGCCGGAACCTGCGGATGCT 4020
405 CAGATTATGACACCGGCTGCGGCTTGAATATTACGACGAGAGACGGGTATCTTGGC 4080
4081 CAGATTATGACACCGGCTGCGGCTTGAATATTACGACGAGAGACGGGTATCTTGGC 4080
465 TGGATGCCGGAAGAAATGACATGATATCCCGTGATTAACCGGCGGCGGCTCGTTC 524
4141 TGGATGCCGGAAGAAATGACATGATATCCCGTGATTAACCGGCGGCGGCTCGTTC 524
525 ATTACGCTTTTGAACCGGCTGAGAGACGGGACAGCTCGGCTGCAATATGCTTTACAG 4200
4201 ATTACGCTTTTGAACCGGCTGAGAGACGGGACAGCTCGGCTGCAATATGCTTTACAG 4200
585 GTGATGAGACAGATGAAGATGCTCGACAGGCTGGAAGAAC 625
4261 GTGATGAGACAGATGAAGATGCTCGACAGGCTGGAAGAAC 4301

```

RESULT 13
 AF434932 6182 bp DNA circular SYN 04-APR-2002
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. 6182
 /organism="Expression vector 410-REV"
 /mol_type="genomic DNA"
 /db_xref="taxon:176974"
 /note="insect-specific"
 97..614
 /note="Autographa californica nuclear polyhedrosis virus
 h75 enhancer"
 615..1201
 /note="Autographa californica nuclear polyhedrosis virus
 immediate early promoter"
 1328..2980
 /codon_start=1
 /product="firefly luciferase"
 /protein_id="AAJ30796.1"
 /db_xref="GI:16904180"
 /translation="MEDAKNIKGPAPFPYPLEDGTAGEOLHKAMKRYALVPGTIAFTD
 AHENVTITAEYFEWVSRLAEAMKRYGLANTNHRIVUSSENSLOFEMFVGALEIVAV
 APANDIYNRELNSMNIISOPTVVFVSKGOKLIVNOKPLTIQKIIIMDSKTVOG
 FOSWTFVTSHLPGFNEVDFPESPRDKTIALIMSSGSGTLPGVALPRTACVR
 FSHARDPIFGNOIIPDTAILSVPHHGFEMFTLGYLICGFVVLMTAFEBELPLRS
 LQDYKIQSALVPTLFSFPASTIIDLKYLDSNHEIASGAPLSKEVGAVKRPHLP
 GIRQYGLTETTSAILITPESGDRPGAVGVVPEFAKVVDLDTGKTIGVNRGELCV
 RGPIMSGYVNNPRATNALIDKQWMLSGDIAVWDEDEHFFIVDLKSLIKKGYOVA
 PAELSTILLOPHNIFDAGVAGLPDDAGELPAVVVLEHGKTMKEKELVIVVASQVTT
 AKKIRGAVVFVDEVPKGLTGKLDARKIREILIKAKKGSKL"
 complement(3166..3298)
 /note="artificial intron"
 3359..3464
 /note="Heliothis virescens p63 chaparonin gene double
 poly-adenylation sequence"
 4381..5241
 /codon_start=1
 /product="beta-lactamase"
 /protein_id="AAJ30797.1"
 /db_xref="GI:16904181"
 /translation="MSIQHFRVALIPPPAFCPLVPNHPETLVYVKDAEDQLGARVGY
 YSPVTEKHLTDGMYRELCSAIIWTSNDTANILITLIGPKELTAFILNMGDHYRL
 DRWEPELNEAIPIIDREDITMPVMAATTLRIKLITGELLTLASRQQLIDWMEADKVAEPL
 EIGASLIKHM"
 5529..5948
 /note="ColE1"

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
 Best Local Similarity 99.3%; Pred. No. 1.7e-145;
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

FEATURES
 source
 1. 6182
 /organism="Expression vector 410-REV"
 /mol_type="genomic DNA"
 /db_xref="taxon:176974"
 /note="insect-specific"
 97..614
 /note="Autographa californica nuclear polyhedrosis virus
 h75 enhancer"
 615..1201
 /note="Autographa californica nuclear polyhedrosis virus
 immediate early promoter"
 1328..2980
 /codon_start=1
 /product="firefly luciferase"
 /protein_id="AAJ30796.1"
 /db_xref="GI:16904180"
 /translation="MEDAKNIKGPAPFPYPLEDGTAGEOLHKAMKRYALVPGTIAFTD
 AHENVTITAEYFEWVSRLAEAMKRYGLANTNHRIVUSSENSLOFEMFVGALEIVAV
 APANDIYNRELNSMNIISOPTVVFVSKGOKLIVNOKPLTIQKIIIMDSKTVOG
 FOSWTFVTSHLPGFNEVDFPESPRDKTIALIMSSGSGTLPGVALPRTACVR
 FSHARDPIFGNOIIPDTAILSVPHHGFEMFTLGYLICGFVVLMTAFEBELPLRS
 LQDYKIQSALVPTLFSFPASTIIDLKYLDSNHEIASGAPLSKEVGAVKRPHLP
 GIRQYGLTETTSAILITPESGDRPGAVGVVPEFAKVVDLDTGKTIGVNRGELCV
 RGPIMSGYVNNPRATNALIDKQWMLSGDIAVWDEDEHFFIVDLKSLIKKGYOVA
 PAELSTILLOPHNIFDAGVAGLPDDAGELPAVVVLEHGKTMKEKELVIVVASQVTT
 AKKIRGAVVFVDEVPKGLTGKLDARKIREILIKAKKGSKL"
 complement(3166..3298)
 /note="artificial intron"
 3359..3464
 /note="Heliothis virescens p63 chaparonin gene double
 poly-adenylation sequence"
 4381..5241
 /codon_start=1
 /product="beta-lactamase"
 /protein_id="AAJ30797.1"
 /db_xref="GI:16904181"
 /translation="MSIQHFRVALIPPPAFCPLVPNHPETLVYVKDAEDQLGARVGY
 YSPVTEKHLTDGMYRELCSAIIWTSNDTANILITLIGPKELTAFILNMGDHYRL
 DRWEPELNEAIPIIDREDITMPVMAATTLRIKLITGELLTLASRQQLIDWMEADKVAEPL
 EIGASLIKHM"
 5529..5948
 /note="ColE1"

every Match 81.3%; Score 574.6; DB 12; Length 7634;

Search completed: January 26, 2005, 08:24:55
Job time : 3017.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:11:40 ; Search time 2221.98 Seconds
(without alignments)
11594.581 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 1 ggcaccatgcgcacattt.....aaactgacgatgggattcc 707

Sequence: 1 ggcaccatgcgcacattt.....aaactgacgatgggattcc 707

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	585.4	82.8	682	CK781302 UI-M-G10-
2	585.4	82.8	703	CF537771 UI-M-G10-
3	585.4	82.8	711	CF743678 UI-M-G10-
4	585.4	82.8	853	CR065566 Forward s
5	584.4	82.7	703	B0154655 NF087C061
6	584.4	82.7	719	CD350897 UI-M-G10-
7	583.8	82.6	751	CR087413 Forward s
8	580.4	82.1	716	CR131675 Forward s
9	580.4	82.0	786	CR108810 Forward s
10	565.4	80.0	691	CD350776 UI-M-G10-
11	564.4	79.8	762	CR035207 Reverse s
12	563.4	79.7	749	CD351273 UI-M-G10-
13	562	79.5	733	CR139473 Forward s
14	555	78.5	633	CR167678 Reverse s
15	554.4	78.4	791	CB520716 UI-M-G10-
16	550.4	77.9	630	CR013093 Reverse s
17	549	77.7	702	CR084987 Reverse s
18	545.4	77.1	611	BT423083 EST533749
19	543.4	76.9	716	CK781284 UI-M-G10-
20	540.8	76.5	716	CF851544 PBMA007XC
21	537.2	76.0	706	AV731514 AV731514
22	535.4	75.7	730	CR077673 Reverse s
23	525	74.3	635	BX982794 Forward s
24	521.8	73.8	536	CR026633 Reverse s

C 25	516	73.0	887	9	CR192748	CR192748 Forward s
C 26	515	72.8	885	9	CR255010	CR255010 Forward s
C 27	498.4	70.5	683	6	CD348896	CD348896 UI-M-FY0-
C 28	493	69.7	495	4	BG457760	BG457760 NF036E03P
C 29	490.4	69.4	786	6	CB520098	CB520098 UI-M-G10-
C 30	482.6	68.3	795	6	CD351738	CD351738 UI-M-G10-
C 31	477.4	67.5	760	7	CF538326	CF538326 UI-M-G10-
C 32	474.4	67.1	620	9	CR034585	CR034585 Forward s
C 33	474.4	67.1	620	9	CR086282	CR086282 Reverse s
C 34	474.4	67.1	620	9	CR171547	CR171547 Reverse s
C 35	471.8	66.7	620	9	CR078862	CR078862 Forward s
C 36	467.4	66.1	752	9	CR088896	CR088896 Forward s
C 37	459.6	65.0	805	6	CD352530	CD352530 UI-M-G10-
C 38	457	64.6	676	5	B0154673	B0154673 NF095C11I
C 39	455.4	64.4	483	9	CR068828	CR068828 Reverse s
C 40	447.4	63.3	547	9	CR148906	CR148906 Forward s
C 41	447.4	63.3	885	9	CR024214	CR024214 Reverse s
C 42	444.4	62.9	599	9	CR079016	CR079016 Forward s
C 43	443.4	62.7	702	9	CR038019	CR038019 Forward s
C 44	440.8	62.3	538	9	CR029330	CR029330 Forward s
C 45	437.6	61.9	901	6	CA328434	CA328434 UI-M-FY0-

ALIGNMENTS

RESULT 1
CK781302/c 682 bp mRNA linear EST 23-FEB-2004
LOCUS UI-M-G10-c1p-g-05-0-UI.r1 NIH-BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:30619060 5', mRNA sequence.
CK781302
VERSION CK781302.1 GI:42746980
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 682)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strauberg, Ph.D.
Email: ggapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Sed primer: PYX-5.
Location/Qualifiers
1..682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone_id="IMAGE:30619060"
/issue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP G10"
/notes="Organ: Brain; Vector: PYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldi, Lemon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 82.8%; Score 585.4; DB 7; Length 682;
Best Local Similarity 99.8%; Pred. No. 2.3e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 45 ATCTAGCTGCATGAGATCATATCGTGGGCTCTTTTCCGGCTAGTCATGCGCCAAAGC 104
Db 611 ATTCAGCTGCATGAGATCATATCGTGGGCTCTTTTCCGGCTAGTCATGCGCCAAAGC 552
Qy 105 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGGTTGAAGCG 164
Db 551 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGGTTGAAGCG 492
Qy 165 GCATGAAAGAGTTTCCCGAGAGATGCTGCTGCAATGACCTTGAGCAAAACGACG 224
Db 491 GCATGAAAGAGTTTCCCGAGAGATGCTGCTGCAATGACCTTGAGCAAAACGACG 432
Qy 225 TTTACCATGATGATTCGGGAAGTGTGCCCATGCAAGCTTTAAAGGTGAAGTGGT 284
Db 431 TTTACCATGATGATTCGGGAAGTGTGCCCATGCAAGCTTTAAAGGTGAAGTGGT 372
Qy 285 CAGGCGACCTGGATTAACAGTTCTGCTGCGGCTTTTCCGAGCAACAGTTCGGATGTCAGC 344
Db 371 CAGGCGACCTGGATTAACAGTTCTGCTGCGGCTTTTCCGAGCAACAGTTCGGATGTCAGC 312
Qy 345 CCGAAGCGCATCAGCAACCCGAAACATTAACCGGAGACGCGAACTGCGCGGTGTG 404
Db 311 CCGAAGCGCATCAGCAACCCGAAACATTAACCGGAGACGCGAACTGCGCGGTGTG 252
Qy 405 CAGATTATGACAGCGGTGCGGCTGGAGATTAACGTACAGAGAGAGGATATCTGGC 464
Db 251 CAGATTATGACAGCGGTGCGGCTGGAGATTAACGTACAGAGAGAGGATATCTGGC 192
Qy 465 TGAATCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTGTC 524
Db 191 TGAATCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTGTC 132
Qy 525 ATTCAAGTTTGAACCCGTGAGAGAGCGGAGACCTGCGGTCGAAATGTTTACAGC 584
Db 131 ATTCAAGTTTGAACCCGTGAGAGAGCGGAGACCTGCGGTCGAAATGTTTACAGC 72
Qy 585 GTGATGAGACAGATGATGATGCTGACACCGCTGAGAAACGCACT 631
Db 71 GTGATGAGACAGATGATGATGCTGACACCGCTGAGAAACGCACT 25

```

RESULT 2
CF537771/c 703 bp mRNA linear EST 12-SEP-2003
LOCUS UI-M-GIO-c1m-m-08-0-UI.r1 NIH BMAP_GIO Mus musculus cDNA clone
DEFINITION IMAGE:30537799 5', mRNA sequence.
ACCESSION CF537771
VERSION CF537771.1 GI:34589753
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 703)
TITLE NIH-MGC <http://mgi.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Jin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
location/Qualifiers
1..703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30537799"
/issue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
/lab_host="PH108 (71 phage resistant)"
/clone_lib="NIH BMAP GIO"
/note="Organ: Brain; Vector: pyx-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
size selected according to RNA size fraction. Ligated
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
program coordinator."

FEATURES

source

ORIGIN

Query Match 82.8%; Score 585.4; DB 7; Length 703;
Best Local Similarity 99.8%; Pred. No. 2.3e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 45 ATCTAGCTGCATGAGATCATATCGTGGGCTCTTTTCCGGCTAGTCATGCGCCAAAGC 104
Db 643 ATTCAGCTGCATGAGATCATATCGTGGGCTCTTTTCCGGCTAGTCATGCGCCAAAGC 584
Qy 105 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGGTTGAAGCG 164
Db 583 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGGTTGAAGCG 524
Qy 165 GCATGAAAGAGTTTCCCGAGAGATGCTGCTGCAATGACCTTGAGCAAAACGACG 224
Db 523 GCATGAAAGAGTTTCCCGAGAGATGCTGCTGCAATGACCTTGAGCAAAACGACG 464
Qy 225 TTTACCATGATGATTCGGGAAGTGTGCCCATGCAAGCTTTTAAAGGTGAAGTGGT 284
Db 463 TTTACCATGATGATTCGGGAAGTGTGCCCATGCAAGCTTTTAAAGGTGAAGTGGT 404
Qy 285 CAGGCGACCTGGATTAACAGTTCTGCTGCGGCTTTTCCGAGCAACAGTTCGATGTCAGC 344
Db 403 CAGGCGACCTGGATTAACAGTTCTGCTGCGGCTTTTCCGAGCAACAGTTCGATGTCAGC 344
Qy 345 CCGAAGCGCATCAGCAACCCGAAACATTAACCGGAGACGCGAACTGCGCGGTGTG 404
Db 343 CCGAAGCGCATCAGCAACCCGAAACATTAACCGGAGACGCGAACTGCGCGGTGTG 284
Qy 405 CAGATTATGACAGCGGTGCGGCTGGAGATTAACGTACAGAGAGAGGATATCTGGC 464
Db 283 CAGATTATGACAGCGGTGCGGCTGGAGATTAACGTACAGAGAGAGGATATCTGGC 224
Qy 465 TGAATCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTGTC 524
Db 223 TGAATCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTGTC 164
Qy 525 ATTCAAGTTTGAACCCGTGAGAGAGCGGAGACCTGCGGTCGAAATGTTTACAGC 584

```

```

Db      163 ATTCACGTTTTTGAAACCCGTCGAGAGAGCGGACACTCGCGGTGCAATGTGTTTTACAGC 104
Qy      585 GTGATGAGCAGATGAGATGCTGCACACGCTGCAGAAACAGCAGCT 631
Db      103 GTGATGAGCAGATGAGATGCTGCACACGCTGCAGAAACAGCAGCT 57

RESULT 3
CF743678/c 711 bp mRNA linear EST 10-OCT-2003
LOCUS      UI-M-G10-clg-e-19-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:30616722 5', mRNA sequence.
ACCESSION  CF743678
VERSION     CF743678.1 GI:37640017
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 711)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mouse1.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1..711
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30616722"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_G10"
/note="Organ: Brain; Vector: PYX-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match      82.8%; Score 585.4; DB 7; Length 711;
Best Local Similarity 99.8%; Pred. No. 2,36-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      45 ATTCAGCTGATCGATGATCATATGCTGCGGCTTTTCCGCTCAGTCATCGCCAGC 104
Db      654 ATTCAGCTGATCGATGATCATATGCTGCGGCTTTTCCGCTCAGTCATCGCCAGC 595
Qy      105 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCCGAGAGTTGAAGC 164
Db      594 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCCGAGAGTTGAAGC 535

```

```

Qy      165 GCATGAAAGAGTTTCCCGAGAGATGACTGCTGCACTTGAAGTGAACGAAACGACG 224
Db      534 GCATGAAAGAGTTTCCCGAGAGATGACTGCTGCTGCACTTGAAGTGAACGAAACGACG 475
Qy      225 TTTCAGATGATGATGAGGAGAGGTCGATGACGACGCTTTTAAAGGTAAGTGTGGT 284
Db      474 TTTCAGATGATGATGAGGAGAGGTCGATGACGACGCTTTTAAAGGTAAGTGTGGT 415
Qy      285 CAGGCCACTGGATATACAGTTTCGTCGCGCTTTTCCGACACAGTTCCGATGTCAGC 344
Db      414 CAGGCCACTGGATATACAGTTTCGTCGCGCTTTTCCGACACAGTTCCGATGTCAGC 355
Qy      345 CCGAAGCGCATACGAAACCCGAAACATACCGGCGACACCGGAACTCGCGTGGTG 404
Db      354 CCGAAGCGCATACGAAACCCGAAACATACCGGCGACACCGGAACTCGCGTGGTG 295
Qy      405 CAGATTATGACAGCGGTGGCGGCTGGGATATTAGCTACGCGAGACGCGGTATCTGGC 464
Db      294 CAGATTATGACAGCGGTGGCGGCTGGGATATTAGCTACGCGAGACGCGGTATCTGGC 235
Qy      465 TGGATGCCGAGAAATGAGATGATACCCGTCGATTAACCCGCGGCGCGCTCGTTC 524
Db      234 TGGATGCCGAGAAATGAGATGATACCCGTCGATTAACCCGCGGCGCGCTCGTTC 175
Qy      525 ATTACGTTTTTTGAACCCGTCGAGAGACGCGGACACTCGCGTCAATGTGTTTACAGC 584
Db      174 ATTACGTTTTTTGAACCCGTCGAGAGACGCGGACACTCGCGTCAATGTGTTTACAGC 115
Qy      585 GTGATGAGCAGATGAGATGCTGCACAGCTGAGAAACGCGAGCT 631
Db      114 GTGATGAGCAGATGAGATGCTGCACAGCTGAGAAACGCGAGCT 68

RESULT 4
CRO65566
LOCUS      Forward strand read from insert in 3'HRT insertion targeting and
DEFINITION chromosome engineering clone MHP250K12, genomic survey sequence.
ACCESSION  CRO65566
VERSION     CRO65566.1 GI:49799156
KEYWORDS    GSS: genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 853)
AUTHORS     Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Rogers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Jonkers,J. and Bradley,A.
TITLE       Direct Submision
JOURNAL     Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            C10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES    Location/Qualifiers
source      1..853
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHP250K12"
            /clone_1lb="MHP2"

ORIGIN
Query Match      82.8%; Score 585.4; DB 9; Length 853;
Best Local Similarity 99.8%; Pred. No. 2,4e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      45 ATTCAGCTGATGATGATCATATGCTGCGGCTTTTCCGCTCAGTCATCGCCAGC 104
Db      264 ATTCAGCTGATGATGATCATATGCTGCGGCTTTTCCGCTCAGTCATCGCCAGC 323
Qy      105 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCCGAGAGTTGAAGC 164
Db      324 TGGCGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCCGAGAGTTGAAGC 383

```

Qy 165 GCATGGAAGAGTTTCCGAGATGACTGTGCTGATGACGTTGAGCCGAAAGCCACG 224
Db 384 GCATGGAAGAGTTTCCGAGATGACTGTGCTGATGACGTTGAGCCGAAAGCCACG 443
Qy 225 TTTCACATGATGATTCGGAAGAGTGGCCATGACAGCCCTTTAACGCTGAACTGTTGCTT 284
Db 444 TTTCACATGATGATTCGGAAGAGTGGCCATGACAGCCCTTTAACGCTGAACTGTTGCTT 503
Qy 285 CAGCCACCTGGGATACCAATGTTGTCGCGCTTTTCCGACACAGTTCCGATGCTGACG 344
Db 504 CAGCCACCTGGGATACCAATGTTGTCGCGCTTTTCCGACACAGTTCCGATGCTGACG 363
Qy 345 CCGAAGCGATCAGCAACCGGAAACAATACCGGCGACAGCGGAACTGCGCTGCTG 404
Db 564 CCGAAGCGATCAGCAACCGGAAACAATACCGGCGACAGCGGAACTGCGCTGCTG 623
Qy 405 CAGATTAAATGACAGCGGTCGCGCTGAGATTAATGCTCAGCAGAGACGCGGATCTGCG 464
Db 624 CAGATTAAATGACAGCGGTCGCGCTGAGATTAATGCTCAGCAGAGACGCGGATCTGCG 683
Qy 465 TGGATGCGCAGAAATGACATGATACCCGCTGATTAACCGGCGGCGCGCTGCTG 524
Db 684 TGGATGCGCAGAAATGACATGATACCCGCTGATTAACCGGCGGCGCGCTGCTG 743
Qy 525 ATTCACGTTTTTGAACCCGCTGAGACGCGGCACTCGCGTGCATATGTTTAAACG 584
Db 744 ATTCACGTTTTTGAACCCGCTGAGACGCGGCACTCGCGTGCATATGTTTAAACG 803
Qy 585 GTGATGAGCAGATGAAGATGCTGACACGCTGCAACACGACGCT 631
Db 804 GTGATGAGCAGATGAAGATGCTGACACGCTGCAACACGACGCT 850

RESULT 5
B0154655/c 703 bp mRNA linear EST 24-APR-2002
LOCUS NF087C06R1P050 Irradiated Medicago truncatula cDNA clone
DEFINITION B0154655
VERSION B0154655.1 GI:20291714
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 703)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W., and May, G.D.,
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula irradiated library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 703 Std Error: 0.00
Plate: 087 row: C column: 06
Seq primer: TCACACAGAAACAGCTATGAC.
FEATURES
Source
Location/Qualifiers
1..703
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF087C06R1P050"
/cissue="NF087C06R1P050"
/dev_stage="seedling"
/clone_id="Irradiated"

/note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m² UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using EXA81st helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOUR cells."

Query Match 82.7%; Score 584.4; DB 5; Length 703;
Best Local Similarity 99.7%; Pred. No. 4,6e-16d;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 45 ATTCAGTCTGATCAGATGATATGCTCGGCTTTTCCGCTCAGTCAATGCGCCAAAGC 104
Db 624 ATTCAGTCTGATCAGATGATATGCTCGGCTTTTCCGCTCAGTCAATGCGCCAAAGC 565
Qy 105 TGGCGGTATCTGGGATGCGGAGAGAGAGCCGCTTTTCCGCGGATTTGAAGCG 164
Db 564 TGGCGGTATCTGGGATGCGGAGAGAGAGCCGCTTTTCCGCGGATTTGAAGCG 505
Qy 165 GCATGGAAGAGTTTCCGAGATGATGCTGCTGCAATGACGTTAGAGGAAAGCAGC 224
Db 504 GCATGGAAGAGTTTCCGAGATGATGCTGCTGCAATGACGTTAGAGGAAAGCAGC 445
Qy 225 TTTCACATGATGATTCGGAAGAGTGGCCATGACAGCCCTTTAACGCTGAACTGTTGCTT 284
Db 444 TTTCACATGATGATTCGGAAGAGTGGCCATGACAGCCCTTTAACGCTGAACTGTTGCTT 385
Qy 285 CAGCCACCTGGGATACCAATGTTGTCGCGCTTTTCCGACACAGTTCCGATGCTGACG 344
Db 384 CAGCCACCTGGGATACCAATGTTGTCGCGCTTTTCCGACACAGTTCCGATGCTGACG 325
Qy 345 CCGAAGCGATCAGCAACCGGAAACAATACCGGCGACAGCCGGAACGCGGCGCGGTG 404
Db 324 CCGAAGCGATCAGCAACCGGAAACAATACCGGCGACAGCCGGAACGCGGCGCGGTG 265
Qy 405 CAGATTAAATGACAGCGGTCGCGCTGAGATTAACCGGCGAGAGCGGATATCTGCG 464
Db 264 CAGATTAAATGACAGCGGTCGCGCTGAGATTAACCGGCGAGAGCGGATATCTGCG 205
Qy 465 TGGATGCGCAGAAATGACATGATACCCGCTGAGTTACCGGCGGCGCGCTGCTG 524
Db 204 TGGATGCGCAGAAATGACATGATACCCGCTGAGTTACCGGCGGCGCGCTGCTG 145
Qy 525 ATTCACGTTTTTGAACCCGCTGAGACGCGGCACTCGGCTGCAATGTTTAAACG 584
Db 144 ATTCACGTTTTTGAACCCGCTGAGACGCGGCACTCGGCTGCAATGTTTAAACG 85
Qy 585 GTGATGAGCAGATGAAGATGCTGACACGCTGCAACACGACGCT 631
Db 84 GTGATGAGCAGATGAAGATGCTGACACGCTGCAACACGACGCT 38

RESULT 6
CD350897/c 719 bp mRNA linear EST 09-JUL-2003
LOCUS CD350897
DEFINITION UI-M-G10-cgh-a-18-0-UI.r1 NIH_BMAP_g10 Mus musculus cDNA clone
IMAGE:6853243 5', mRNA sequence.
ACCESSION CD350897
VERSION CD350897.1 GI:31142412
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 719)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-rc@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES

Seq primer: PYX-5.
 Location/Qualifiers

```

1..719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853243"
/libase_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="MD10B (T1 phage resistant)"
/clone_lib="NIH BMAP G10"
/note="Organ: Brain; Vector: PYX-Asc; Site_1: Ecor I; Site_2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the university Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

```

Query Match      82.7%; Score 584.4; DB 6; Length 719;
Best Local Similarity 99.7%; Pred. No. 4.6e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGATCATATCGTCGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 663 ATCCAGCTGCATCAGATCATATCGTCGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 604
QY 105 TGGGCTATCTGGGCATCGGGAGAGAAAGCCCGTGCCTTTTTCGGGAGGTTGAAGCG 164
DB 603 TGGGCTATCTGGGCATCGGGAGAGAAAGCCCGTGCCTTTTTCGGGAGGTTGAAGCG 544
QY 165 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTCAGCTTGAAGCAAAACGACG 224
DB 543 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTCAGCTTGAAGCAAAACGACG 484
QY 225 TTTCACATGATGATTCGGGAAAGTGTGGCCATGACGCGCTTTTAAACGATGAACTGTTGTT 284
DB 483 TTTCACATGATGATTCGGGAAAGTGTGGCCATGACGCGCTTTTAAACGATGAACTGTTGTT 424
QY 285 CAGGCCACTTGGGATACCAAGTTGTCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 344
DB 423 CAGGCCACTTGGGATACCAAGTTGTCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 364
QY 345 CCGAAGGCAATCCGGAACCAATACCGGCGACACGCGGAACTGCGGTCGGTGTG 404
DB 363 CCGAAGGCAATCCGGAACCAATACCGGCGACACGCGGAACTGCGGTCGGTGTG 304
QY 405 CAGATTATGACAGCGGTGGCGGCTGGGATATTACGTCAGCGGATCTCTGGC 464

```

```

DB 303 CAGATTATGACAGCGGTGGCGGCTGGGATATTACGTCAGCGGATCTCTGGC 244
QY 465 TGAATGCCGCAAAATGACATGATACCCCGTAGTTACCCGCGGCGCGCTCGTTG 524
DB 243 TGAATGCCGCAAAATGACATGATACCCCGTAGTTACCCGCGGCGCGCTCGTTG 184
QY 525 ATTACGCTTTTGAACCCGTCGAGAGAGGAGAGCTGCGGCTGGAATGTCCTTTTACAGC 584
DB 183 ATTACGCTTTTGAACCCGTCGAGAGAGGAGAGCTGCGGCTGGAATGTCCTTTTACAGC 124
QY 585 GTGATGAGCAGATGATGATGCTGACACGCTGCGAACAACGACAGCT 631
DB 123 GTGATGAGCAGATGATGATGCTGACACGCTGCGAACAACGACAGCT 77

```

RESULT 7
 CR087413/c 751 bp DNA linear GSS 05-JUL-2004
 LOCUS CR087413
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP438f24, genomic survey sequence.
 ACCESSION CR087413
 VERSION CR087413.1 GI:49821005
 KEYWORDS GSS; genome survey sequence; MGCER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus. 1 (bases 1 to 751)
 AUTHORS Adams,D.J., Biggs,J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>

FEATURES

source
 Location/Qualifiers
 1..751
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP438f24"
 /clone_lib="MHP"

ORIGIN

```

Query Match      82.6%; Score 583.8; DB 9; Length 751;
Best Local Similarity 99.7%; Pred. No. 7e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGATCATATCGTCGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 623 ATCCAGCTGCATCAGATCATATCGTCGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 564
QY 105 TGGGCTATCTGGGCATCGGGAGAGAAAGCCCGTGCCTTTTTCGGGAGGTTGAAGCG 164
DB 563 TGGGCTATCTGGGCATCGGGAGAGAAAGCCCGTGCCTTTTTCGGGAGGTTGAAGCG 504
QY 165 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTCAGCTTGAAGCAAAACGACG 224
DB 503 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTCAGCTTGAAGCAAAACGACG 444
QY 225 TTTCACATGATGATTCGGGAAAGTGTGGCCATGACGCGCTTTTAAACGATGAACTGTTGTT 284
DB 443 TTTCACATGATGATTCGGGAAAGTGTGGCCATGACGCGCTTTTAAACGATGAACTGTTGTT 384
QY 285 CAGGCCACTTGGGATACCAAGTTGTCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 344
DB 383 CAGGCCACTTGGGATACCAAGTTGTCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 324
QY 345 CCGAAGGCAATCCGGAACCAATACCGGCGACACGCGGAACTGCGGTCGGTGTG 404
DB 323 CCGAAGGCAATCCGGAACCAATACCGGCGACACGCGGAACTGCGGTCGGTGTG 264
QY 405 CAGATTATGACAGCGGTGGCGGCTGGGATATTACGTCAGCGGATCTCTGGC 464

```

Db 263 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGCGGTATCTTGGC 204
QY 465 TGGATGCCGCGAAGAAATGACATGATATACCCCGTGTATACCCGGGGGCGCGCTCTTGC 204
Db 203 TGGATGCCGCGAAGAAATGACATGATATACCCCGTGTATACCCGGGGGCGCGCTCTTGC 524
QY 525 ATTACAGTTTTTGAACCCGTGAGAGACCGGAGACTCGCGGTGCAAAATGTGTTTACAGC 144
Db 143 ATTACAGTTTTTGAACCCGTGAGAGACCGGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
QY 585 GTGATGAGCAGATGAAATGCTGACACCGCTGACAAACAGGAGCT 631
Db 83 GTGATGAGCAGATGAAATGCTGACACCGCTGACAAACAGGAGCT 37

RESULT 8
CRI31675/c
LOCUS
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
ACCESSION CRI31675.1 GI:49879128
VERSION CRI31675.1 GI:49879128
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 716)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR
location/Qualifiers
source 1..716
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP438f23"
/clone_1lb="MHP"

ORIGIN
Query Match 82.1%; Score 580.4; DB 9; Length 716;
Best Local Similarity 99.8%; Pred. No. 7.3e-163;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATTACCTGATCAGATCATATGTCGGGTCTTTTCCGCTCAGTCAATGCCCAAGC 104
Db 582 ATTCAGCTGATCAGATCATATGTCGGGTCTTTTCCGCTCAGTCAATGCCCAAGC 104
QY 105 TGGCGCTATCTGGGAGATGGGAGAAAGCCCGTGTCTTTCCCGGAGATTGAACG 523
Db 522 TGGCGCTATCTGGGAGATGGGAGAAAGCCCGTGTCTTTCCCGGAGATTGAACG 523
QY 165 GATGAGAAAGATTGGCGAGATGACTGCTGTGATGATGAGCTTGAACGAAACGACG 463
Db 462 GATGAGAAAGATTGGCGAGATGACTGCTGTGATGATGAGCTTGAACGAAACGACG 463
QY 225 TTTCACATGATGATGGGAGAGGTGGCCATGACGCTTTAACGAGTGAATCTTGTG 284
Db 402 TTTCACATGATGATGGGAGAGGTGGCCATGACGCTTTAACGAGTGAATCTTGTG 284
QY 285 CAGGCCACCTGGGATACCAATTCGTGCGGCTTTTCCGAGACAGTTCCGATGTCAAC 343
Db 342 CAGGCCACCTGGGATACCAATTCGTGCGGCTTTTCCGAGACAGTTCCGATGTCAAC 343
QY 345 CCGAAGCGCATCAGCAACCGAACAATTCGGCGACAGCCGGAATCTGCGGTGTG 283
Db 282 CCGAAGCGCATCAGCAACCGAACAATTCGGCGACAGCCGGAATCTGCGGTGTG 283
QY 405 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGCGGTATCTTGGC 464
Db 464 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGCGGTATCTTGGC 464

Db 222 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGCGGTATCTTGGC 163
QY 465 TGGATGCCGCGAAGAAATGACATGATATACCCCGTGTATACCCGGGGGCGCGCTCTTGC 163
Db 162 TGGATGCCGCGAAGAAATGACATGATATACCCCGTGTATACCCGGGGGCGCGCTCTTGC 524
QY 525 ATTACAGTTTTTGAACCCGTGAGAGACCGGAGACTCGCGGTGCAAAATGTGTTTACAGC 103
Db 102 ATTACAGTTTTTGAACCCGTGAGAGACCGGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
QY 585 GTGATGAGCAGATGAAATGCTGACACCGCTGACAAACAGGAGCT 626
Db 42 GTGATGAGCAGATGAAATGCTGACACCGCTGACAAACAGGAGCT 43

RESULT 9
CRI08810/c
LOCUS
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
ACCESSION CRI08810
VERSION CRI08810.1 GI:49856225
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 786)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR
location/Qualifiers
source 1..786
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP22016"
/clone_1lb="MHP"

ORIGIN
Query Match 82.0%; Score 580; DB 9; Length 786;
Best Local Similarity 100.0%; Pred. No. 9.8e-163;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TGCATCAGATCATATGTCGGGTCTTTTCCGCTCAGTCAATGCCCAAGCTGGCGCT 111
Db 786 TGCATCAGATCATATGTCGGGTCTTTTCCGCTCAGTCAATGCCCAAGCTGGCGCT 111
QY 112 ATTCGGGATCTGGGAGAGAAAGACCCGTCCTTTTCCCGGAGATTGAACGCGCATGGA 171
Db 726 ATTCGGGATCTGGGAGAGAAAGACCCGTCCTTTTCCCGGAGATTGAACGCGCATGGA 171
QY 172 AAGGTTTGGCGAGATGACTGCTGTGATGATGAGCTTGAACGAAACGACGTTTACCA 231
Db 666 AAGGTTTGGCGAGATGACTGCTGTGATGATGAGCTTGAACGAAACGACGTTTACCA 231
QY 232 TGAATGATGGGAGAGGTGGCCATGACAGCTTTTAAAGTGAACGTTGCTGTTACAGCCA 291
Db 606 TGAATGATGGGAGAGGTGGCCATGACAGCTTTTAAAGTGAACGTTGCTGTTACAGCCA 291
QY 292 CCGGAGATACCAATTCGTGCGGCTTTTCCGAGACAGTTCCGATGTCAACCGCA 547
Db 546 CCGGAGATACCAATTCGTGCGGCTTTTCCGAGACAGTTCCGATGTCAACCGCA 547
QY 352 GCATCAGCAACCGCAACAATACCGGAGACAGCCGGAATCTGCGGTGTGAGAGTTA 411
Db 486 GCATCAGCAACCGCAACAATACCGGAGACAGCCGGAATCTGCGGTGTGAGAGTTA 411
QY 412 ATGACAGCGGTGGCGCTGGGATATTACGTACAGGAGAGCGGTATCTTGGCTGATGC 471
Db 471 ATGACAGCGGTGGCGCTGGGATATTACGTACAGGAGAGCGGTATCTTGGCTGATGC 471

Db 426 ATGACGCGGTGGCGGCTGGGATATTAAGTCAGGAGAGACGGGATTCCTGGCTGATGTC 367

Qy 472 CGCAGAAATGACATGATATACCCCTGAGTTACCCGGGGGGCGCCCTGCTTCATTACAG 531

Db 366 CGCAGAAATGACATGATATACCCCTGAGTTACCCGGGGGGCGCCCTGCTTCATTACAG 307

Qy 532 TTTTGAACCCGTGGAGAGACGGGACAGACTCGCGGGTGCAGAAATGTGTTTAAAGCCGATGAG 591

Db 306 TTTTGAACCCGTGGAGAGACGGGACAGACTCGCGGGTGCAGAAATGTGTTTAAAGCCGATGAG 247

Qy 592 AGCAGATGAGATGCTCGACACGCTGACAGAACCCGACACT 631

Db 246 AGCAGATGAGATGCTCGACACGCTGACAGAACCCGACACT 207

RESULT 10
CD350776/c 691 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-G10-csh-9-19-0-UI_r1 NIH_BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:6853388 5', mRNA sequence.
ACCESSION CD350776 GI:31142363
VERSION CD350776.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 691)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsbbs-remail.nih.gov

Tissue Procurement: Dr. Jim Jin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
source
Seq primer: PYX-5.
Location/Qualifiers

1..691
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853388"
/feature_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP G10"
/note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN
Query Match 80.0%; Score 565.4; DB 6; Length 691;
Best Local Similarity 99.8%; Pred. No. 2,3e-158;
Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGATCATATGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 104

Db 567 ATTCAGCTGCATCAGATCATATGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 508

Qy 105 TGGGCGCTATCTGGGCGATCGGGAGAGAAACCCGTCCTTTCCCGGAGGTTGAAGC 164

Db 507 TGGGCGCTATCTGGGCGATCGGGAGAGAAACCCGTCCTTTCCCGGAGGTTGAAGC 448

Qy 165 GCATGAAAGAGTTTGGCCGAGAGTACTGCTGTCATTGACGTTGACGAAACGACAGC 224

Db 447 GCATGAAAGAGTTTGGCCGAGAGTACTGCTGTCATTGACGTTGACGAAACGACAGC 388

Qy 225 TTACCATGATGATTTCCGGAAGGTGTGCGCATGACGCTTTTAAAGTGAACCTGTCCTT 284

Db 387 TTACCATGATGATTTCCGGAAGGTGTGCGCATGACGCTTTTAAAGTGAACCTGTCCTT 328

Qy 285 CAGGCGACCTGGGATACAGATTGTCGCGGCTTTTCCGGACACAGTTCCGGATGCTGACG 344

Db 327 CAGGCGACCTGGGATACAGATTGTCGCGGCTTTTCCGGACACAGTTCCGGATGCTGACG 268

Qy 345 CCGAAGCGCATCAGCAACCCGACCAATACCGGACAGCCGGAACCTGCGCTGCGGCTGTG 404

Db 267 CCGAAGCGCATCAGCAACCCGACCAATACCGGACAGCCGGAACCTGCGCTGCGGCTGTG 208

Qy 405 CAGATTAATGACACCGCTGCGGCTGGGATATTACGTCACGAGACGCGGTATCTGCGC 464

Db 207 CAGATTAATGACACCGCTGCGGCTGGGATATTACGTCACGAGACGCGGTATCTGCGC 148

Qy 465 TGGATGCCCGGAGAAATGAGCAATGATACCCCGTAGTTACCCGGCGGGCGGCTCGTTC 524

Db 147 TGGATGCCCGGAGAAATGAGCAATGATACCCCGTAGTTACCCGGCGGGCGGCTCGTTC 88

Qy 525 ATTACGCTTTTGAACCCGTGAGAGACGCGGACAGCTCGCGTGAATGTGTTTACAGC 584

Db 87 ATTACGCTTTTGAACCCGTGAGAGACGCGGACAGCTCGCGTGAATGTGTTTACAGC 28

Qy 585 GTGATGAGCAGATGAAGATGCTGAC 611

Db 27 GTGATGAGCAGATGAAGATGCTGAC 1

RESULT 11
LOCUS CR035207 762 bp DNA linear GSS 05-JUL-2004
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN278n13, genomic survey sequence.
ACCESSION CR035207 GI:49768262
VERSION GSS: genome survey sequence; MICE.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 762)
ADAMS,D.J., BIGGS,P.J., COX,A.V., DAVIES,R.M., VAN DER WEYDEN,L., ROGERS,J., SMITH,J., PLUMB,R.W., TAYLOR,R.G., NISHIJIMA,I., YU,Y., ROGERS,J. and Bradley,A.

TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICE

FEATURES
source
Location/Qualifiers
1..762
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN278n13"
/clone_lib="MHPN"

ORIGIN
Query Match 79.8%; Score 564; DB 9; Length 762;
Best Local Similarity 100.0%; Pred. No. 6,2e-158;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 CGTGGGCTCTTTTTCGGGCTCAGTCATGCCCAAGCTGGCGTATCTGGGCATCGGGGA 127
 Db 2 CGTGGGCTCTTTTTCGGGCTCAGTCATGCCCAAGCTGGCGCTATCTGGGCATCGGGGA 127
 QY 128 GGAAGAAGCCCGGCGCTTTTCCCGGAGGTTGAACCGCATGAAAAGTTGCCAGGA 61
 Db 62 GGAAGAAGCCCGGCGCTTTTCCCGGAGGTTGAACCGCATGAAAAGTTGCCAGGA 187
 QY 188 TCAGTCGTCGTCATTTGACGTTGAGCGAAAACGACGTTTACATGATGTTTCGGGAAG 121
 Db 122 TCAGTCGTCGTCATTTGACGTTGAGCGAAAACGACGTTTACATGATGTTTCGGGAAG 247
 QY 248 TGTGGCCATGCAAGCCGCTTTAACGAGTAACTGTTGTTCAAGCCACCTGGAGTACCAAGTTC 307
 Db 182 TGTGGCCATGCAAGCCGCTTTAACGAGTAACTGTTGTTCAAGCCACCTGGAGTACCAAGTTC 307
 QY 308 GTCCGCGCTTTTCCGGACACAGTTCCGAGTGTGACCCGGAAGCGCATACGAAACCCGAA 367
 Db 242 GTCCGCGCTTTTCCGGACACAGTTCCGAGTGTGACCCGGAAGCGCATACGAAACCCGAA 367
 QY 368 CAATTACCGGAGACAGCCGGAACTGCGTCCGTCGCGTGTGAGTTATATGACAGCGTGGCG 427
 Db 302 CAATTACCGGAGACAGCCGGAACTGCGTCCGTCGCGTGTGAGTTATATGACAGCGTGGCG 427
 QY 428 GCTGGAGATTTACGTCAAGAGAGAGGGTATCTCGGCTGATGCGCAGAAATGACATG 361
 Db 362 GCTGGAGATTTACGTCAAGAGAGAGGGTATCTCGGCTGATGCGCAGAAATGACATG 487
 QY 488 GATACCCCGTAGTTTACCAGCGGCGCGGCTCGTTATTCACGTTTTTGAACCCGTGGA 421
 Db 422 GATACCCCGTAGTTTACCAGCGGCGCGGCTCGTTATTCACGTTTTTGAACCCGTGGA 547
 QY 548 GAGCGGGCAGACTGCGCGTCAATGTGTTTTACGCGTATGAGCAGATGATAGATGCT 481
 Db 482 GAGCGGGCAGACTGCGCGTCAATGTGTTTTACGCGTATGAGCAGATGATAGATGCT 607
 QY 608 CGACACCCCTCAGAAACCGCAGCT 631
 Db 542 CGACACCCCTCAGAAACCGCAGCT 565

```

RESULT 12
CD351273/c
LOCUS
DEFINITION      CD351273              749 bp
                  U1-M-G10-cg-h-f-22-0-U1.r1 NIH BMAP mRNA
ACCESSION      IMAGE:6853367 5', mRNA sequence.
VERSION        CD351273
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.iowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seg primer: pyx-5.
Location/Qualifiers
1. 749
/organism="Mus musculus"
/mol_type="mRNA"

```

```

/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="RIMAG:6853367"
/tissue="whole brain"
/dep_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (71 phage resistant)"
/clone_id="N1H-BMAP_G10"
/notes="Organ: Brain; Vector: pX-Asc; Site:1: EcoR I;
Site:2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction. ligated
with EcoR I adaptor digested with NotI and then cloned
directionally into pX-Asc vector. The library tag
is ACCGAGACAGC. This library was created for the polya tail
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
Program coordinator."

```

Query Match	Best Local Similarity	99.8%	Score 563.4	DB 6	Length 749
Matches 564	Conservative	0	Mismatches 1	Indels 0	Gaps
Qy	45	ATCTAGCTGCATCAGATCATATGCTGCTGGGCTCTTTTTCGCGCTCAATCATGCCCCAAGC			
Db	565	ATCTAGCTGCATCAGATCATATGCTGCTGGGCTCTTTTTCGCGCTCAATCATGCCCCAAGC			
Qy	105	TGCGCGTATCTGGGGATTCGGGAGGAAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGC			
Db	505	TGCGCGTATCTGGGATCTGGGAGGAGAAAGAACCCGCTTTTCCGCGAGGTTGAAGC			
Qy	165	GCATGAAAAGGTTTGCCTGAGATGATCTGCTGCTGATGACGTTGAGGAAAAGCAGC			
Db	445	GCATGAAAAGGTTTGCCTGAGATGATCTGCTGCTGATGACGTTGAGGAAAAGCAGC			
Qy	225	TTTACATGATGATTTGGGGAAGGTGTGGCCATGCAAGCCTTTAAAGGAACTGTTGCTT			
Db	385	TTTACATGATGATTTGGGGAAGGTGTGGCCATGCAAGCCTTTAAAGGAACTGTTGCTT			
Qy	285	CAGGCCACCTGGGATACAGCTTCGTGCGGCTTTTCCGGAACACGTTCCGATGGTCAGC			
Db	325	CAGGCCACCTGGGATACAGCTTCGTGCGGCTTTTCCGGAACACGTTCCGATGGTCAGC			
Qy	345	CCGAAGCGCATCAGAACCCGGAACAATTCGCGGCAACGCTGAACTGCGCGGATGTCAG			
Db	265	CCGAAGCGCATCAGAACCCGGAACAATTCGCGGCAACGCTGAACTGCGCGGATGTCAG			
Qy	405	CAGATTAATGACAGCGGTCGCGGCTGGGATATTAGCTGACGAGGACGGATATCTTGAC			
Db	205	CAGATTAATGACAGCGGTCGCGGCTGGGATATTAGCTGACGAGGACGGATATCTTGAC			
Qy	465	TGGAATCCGCGCAAAATGACATGATATACCCCGGAACTTACCAGCGGCGCTCGTTC			
Db	145	TGGAATCCGCGCAAAATGACATGATATACCCCGGAACTTACCAGCGGCGCTCGTTC			
Qy	525	ATTACGCTTTTGAACCCCGTGAAGGACGGGCAAGCTCGCGTGCACAAATGTTTTTAAGC			
Db	85	ATTACGCTTTTGAACCCCGTGAAGGACGGGCAAGCTCGCGTGCACAAATGTTTTTAAGC			
Qy	585	GTCATGGAAGATTAAGATGCTCG 609			
Db	25	GTCATGGAAGATTAAGATGCTCG 1			

Chromosome engineering clone MHP274b23, genomic survey sequence.

ACCESSION CR139473.1 GI:49887345
VERSION CR139473.1
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 733)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.

TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>
Location/Qualifiers

FEATURES
source 1..733
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP274b23"
/clone_1id="MHP"

ORIGIN

Query Match 79.5%; Score 562; DB 9; Length 733;
Best Local Similarity 99.3%; Pred. No. 2.5e-157;
Matches 573; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

56 TCAGGATCATATCTCGGGCTCTTTTCCGGCTCAGTATGCGCCAGACTGGCGTATCT 115
733 TCAGGATCATATCTCGGGCTCTTTTCCGGCTCAGTATGCGCCAGACTGGCGTATCT 674
116 GGGCATCGGGAGAGAAAGCCCGCTCTTTCCCGAGGTTGAAGCGGATGGAAGA 175
673 GGGCATCGGGAGAGAAAGCCCGCTCTTTCCCGAGGTTGAAGCGGATGGAAGA 614
176 GTTTCGCGAGGATACCTGCTGCTGCTGATGAGTGAAGGAAAGGCAAGTTACCATAT 235
613 GTTTCGCGAGGATACCTGCTGCTGCTGATGAGTGAAGGAAAGGCAAGTTACCATAT 554
236 GATTGCGGAAAGTGTGCGCATGCAAGCCTTTAAAGGTAAGTGTGTTGTTGAGGCACTG 295
553 GATTGCGGAAAGTGTGCGCATGCAAGCCTTTAAAGGTAAGTGTGTTGTTGAGGCACTG 494
296 GGATACCAAGTGTGCGCGCTTTTCCGAGACAAGTTCGGATGTGACGCCGGAAGCGCAT 355
493 GGATACCAAGTGTGCGCGCTTTTCCGAGACAAGTTCGGATGTGACGCCGGAAGCGCAT 434
356 CAGGAACCCGGAACAATACCGCGGACAGCGGGAAGTGTGCGGATGTGACGCCGGAAGCGCAT 414
433 CAGGAACCCGGAACAATACCGCGGACAGCGGGAAGTGTGCGGATGTGACGCCGGAAGCGCAT 374
415 ACACGCGTGCAGGCTGAGATATTAACGTACAGAGAGAGGATATCTGGCTGATGAGCGC 474
373 ACACGCGTGCAGGCTGAGATATTAACGTACAGAGAGAGGATATCTGGCTGATGAGCGC 314
475 AGAATGAGACATGATACCCGCTGAGTATACCGCGGAGCGGCTCGTTCACTACGTTT 534
313 AGAATGAGACATGATACCCGCTGAGTATACCGCGGAGCGGCTCGTTCACTACGTTT 254
535 TTGAACCCGTTGAG 594
253 TTGAACCCGTTGAG 194
595 AGATGAGATGCTGACACGCTGAGAAACAGCAGCT 631
193 AGATGAGATGCTGACACGCTGAGAAACAGCAGCT 157

RESULT 14
CR167678 633 bp DNA linear GSS 06-JUN-2004
LOCUS CR167678
DEFINITION Reverse strand read from insert in 3'HPT insertion targeting and

Chromosome engineering clone MHP161K02, genomic survey sequence.

ACCESSION CR167678 GI:49946527
VERSION CR167678.1
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 633)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.

TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>
Location/Qualifiers

FEATURES
source 1..633
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP161K02"
/clone_1id="MHP"

ORIGIN

Query Match 78.5%; Score 555; DB 9; Length 633;
Best Local Similarity 100.0%; Pred. No. 3e-155;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

77 TTTTTCGCGGCTCAGTATGCGCCAGCTGCGCTATCTGGGCAATCGGGAGAGAAGC 136
633 TTTTTCGCGGCTCAGTATGCGCCAGCTGCGCTATCTGGGCAATCGGGAGAGAAGC 574
137 CCGTGCCTTTTCCCGAGGTTGAAGCGGATGGAAGAGTTTCCGAGATGACTGCTG 196
573 CCGTGCCTTTTCCCGAGGTTGAAGCGGATGGAAGAGTTTCCGAGATGACTGCTG 514
197 CTGATTTGAGCTTGAAGGAAAGCAGCTTTTACATATGATTTGGGAAAGTGTGGCCAT 256
513 CTGATTTGAGCTTGAAGGAAAGCAGCTTTTACATATGATTTGGGAAAGTGTGGCCAT 454
257 GCAGCGCTTTTAAAGGTAAGTGTGTTGTTGAGGCACTGGATACCAATTCAGTGTGCGGCT 316
453 GCAGCGCTTTTAAAGGTAAGTGTGTTGTTGAGGCACTGGATACCAATTCAGTGTGCGGCT 394
317 TTTCCGAGACAGTTCGGATGTGACGCCGGAAGCGATGCAAGCCGGAACAATACCGG 376
393 TTTCCGAGACAGTTCGGATGTGACGCCGGAAGCGATGCAAGCCGGAACAATACCGG 334
377 CGACAGCCGGAACAATGCGGCTGCGGATGTCAGATTAAAGACGCGGTGCGGCTGGGATA 436
333 CGACAGCCGGAACAATGCGGCTGCGGATGTCAGATTAAAGACGCGGTGCGGCTGGGATA 274
437 TTACGTCAG 496
273 TTACGTCAG 214
497 TGAATTACCCGCGCGGCGCTGTTCACTTCAAGTTTGAACCCGTTGAGAGAGAGAG 556
213 TGAATTACCCGCGCGGCGCTGTTCACTTCAAGTTTGAACCCGTTGAGAGAGAGAG 154
557 GACTCGGCGTGAAGTGTGTTTACAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 616
153 GACTCGGCGTGAAGTGTGTTTACAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 94
617 GCAGAACCGGAGCT 631
93 GCAGAACCGGAGCT 79

RESULT 15
CB520716 791 bp mRNA linear EST 09-JUN-2003
LOCUS CB520716
DEFINITION UT-M-G10-cej-j-06-0-UT.r1 NIH_BMAP_G10 Mus musculus cDNA clone

IMAGE:6840391.5', mRNA sequence.
 ACCESSION CBS20716 GI:29354071
 VERSION CBS20716.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 791)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 source
 Seq primer: PYX-5.
 Location/Qualifiers
 1..791
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6840391"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="N1H_BMAP_610"
 /note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into PYX-Asc vector. The library tag
 is AGCGAGACG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery' in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 78.4%; Score 554.4; DB 6; Length 791;
 Best Local Similarity 99.6%; Pred. No. 4.8e-155;
 Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 45 ATCTAGCTGCATCAGATCATATGCTGGGCTTTTTCGGGCTCAGTCATGCCCAAGC 104
 DB 568 ATCCAGCTGCATCAGATCATATGCTGGGCTTTTTCGGGCTCAGTCATGCCCAAGC 509
 QY 105 TGGGCTATCTGGGCGATCGGGAGAGAAAGCCGTCCTTTTCCGGAGGTTGAACG 164
 DB 508 TGGGCTATCTGGGCGATCGGGAGAGAAAGCCGTCCTTTTCCGGAGGTTGAACG 449
 QY 165 GCATGGAAGAGTTTCCGAGATGACTGCTGCTGATTTAGCGGAAACGACG 224
 DB 448 GCATGGAAGAGTTTCCGAGATGACTGCTGCTGATTTAGCGGAAACGACG 389
 QY 225 TTACATATGATTTGGGAAGGTGTGGCATGACGCTTTAAACGAGTGTGTT 284
 DB 388 TTACATATGATTTGGGAAGGTGTGGCATGACGCTTTAAACGAGTGTGTT 329
 QY 285 CAGGCGACCTGGGATACGATTGTCGGGGCTTTCCGACACAGTTCGGATGTCAGC 344

DB 328 CAGGCGACCTGGGATACGATTGTCGGGGCTTTTCCGACACAGTTCGGATGTCAGC 269
 QY 345 CCGAAGCCGATCAGCAACCCGACAAATACCGGACAGCCGGAATCGCGTGCCTGTG 404
 DB 268 CCGAAGCCGATCAGCAACCCGACAAATACCGGACAGCCGGAATCGCGTGCCTGTG 209
 QY 405 CAGATTAATGACAGCGGTGGCGCGCTGGGATTTAGTCACGAGACGGTATCCTGAC 464
 DB 208 CAGATTAATGACAGCGGTGGCGCGCTGGGATTTAGTCACGAGACGGTATCCTGAC 149
 QY 465 TGGATGCCGAGAAATGACATGATACCCGTAAGTTACCCGGCGGGCGGCTTCGTT 524
 DB 148 TGGATGCCGAGAAATGACATGATACCCGTAAGTTACCCGGCGGGCGGCTTCGTT 89
 QY 525 ATTACGTTTTTGAACCCGTTGAGAGACGGGACAGACTCGCGGTGCAATGTGTTTAC-AG 583
 DB 88 ATTACGTTTTTGAACCCGTTGAGAGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAG 29
 QY 584 CCGTATGAGACAGATGAAAGATGCTCGAC 611
 DB 28 CGTATGAGACAGATGAAAGATGCTCGAC 1

Search completed: January 26, 2005, 12:13:54
 Job time : 2228.98 secs

to the end of the piggyBac molecule. The DNA molecule in the transposon piggyBac is useful for transferring genes into host cells or embryos for transforming the cells of embryos. The transformed cells or embryos are useful for developing or making transgenic organisms. This sequence represents a minimal sequence cartridge of the eukaryotic transformation vector piggyBac

Sequence 707 BP, 164 A, 174 C, 208 G, 161 T, 0 U, 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 707;
Best Local Similarity 100.0%; Pred. No. 2,3e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGATCCCATGCGTCAATTTTACGAGACTATCTTTTACGGTTATCTAGCTGCATCAGG 60
DB 1 GGATCCCATGCGTCAATTTTACGAGACTATCTTTTACGGTTATCTAGCTGCATCAGG 60
QY 61 ATCATATGCTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGCA 120
DB 61 ATCATATGCTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGCA 120
QY 121 TCGGGAGAGAGAACCCGCTTTTCCCGCAAGTTGAAGCGCATGAAAGATTG 180
DB 121 TCGGGAGAGAGAACCCGCTTTTCCCGCAAGTTGAAGCGCATGAAAGATTG 180
QY 181 CCGAGATGACTGCTGCTGATTCAGCTTGAGCGAAGACGCTTACATGATGATTC 240
DB 181 CCGAGATGACTGCTGCTGATTCAGCTTGAGCGAAGACGCTTACATGATGATTC 240
QY 241 GGGAGAGTGTGGCCATGACAGCCCTTTTAAAGCTGTTGCTTACAGCCAGCTGGATA 300
DB 241 GGGAGAGTGTGGCCATGACAGCCCTTTTAAAGCTGTTGCTTACAGCCAGCTGGATA 300
QY 301 CCGATTCGTGCGCGCTTTTTCGGACACAGTTCCGATGAGTGCAGCCCGAAGGCAATGCA 360
DB 301 CCGATTCGTGCGCGCTTTTTCGGACACAGTTCCGATGAGTGCAGCCCGAAGGCAATGCA 360
QY 361 ACCCGAACAATACCGCGCAGACCGGAACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 ACCCGAACAATACCGCGCAGACCGGAACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 GTGCGCGCTGCGGATTAATGCTGAGCGAGACGGGTATCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GTGCGCGCTGCGGATTAATGCTGAGCGAGACGGGTATCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 GGACATGATACCCCGTGAATTAACCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GGACATGATACCCCGTGAATTAACCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CCGTGGAGAGAGCGGACAGCTGCGGTCAATGTTTACAGCGTGAAGAGAGATGA 600
DB 541 CCGTGGAGAGAGCGGACAGCTGCGGTCAATGTTTACAGCGTGAAGAGAGATGA 600
QY 601 AGATGCTGCAACCGCTGCAAGACAGCTGTAATTAACCTTAAAGAAATTAATCAATTG 660
DB 601 AGATGCTGCAACCGCTGCAAGACAGCTGTAATTAACCTTAAAGAAATTAATCAATTG 660
QY 661 TGAGTACGTTAAAGATTAATCATGCTTAAATTAAGCGATGGGATTC 707
DB 661 TGAGTACGTTAAAGATTAATCATGCTTAAATTAAGCGATGGGATTC 707

```

RESULT 2
ACAS5360
ID ACAS5360 standard; DNA; 4613 BP.

AC ACAS5360;

XX 06-JUN-2003 (first entry)

DE Transformation vector piggyBac related plasmid pCR11-1TR.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;

KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.

XX Synthetic.

PN US2002173634-A1.

PD 21-NOV-2002.

PF 30-OCT-2001; 2001US-00001189.

PR 31-OCT-2000; 2000US-0244677P.

PA 01-NOV-2000; 2000US-0244984P.

PA (FRAS/) FRASER M J.

PA (LIXX/) LI X.

PA (BEAM/) BEAM T.

PA (HUA/) HUA-VAN A.

PI Fraser MJ, Li X, Beam T, Hua-Van A;

DR WPI; 2003-352597/33.

DR P-PSDB; ABUT0357, ABUT0358.

PT New DNA molecule in the transposon piggyBac, useful for transferring

PT genes into host cells or embryos for transforming the cells of embryos

PT that can be used in making transgenic organisms.

PS Disclosure; Fig 10(B); 151pp; English.

CC The invention describes a DNA molecule comprising at least 163

CC consecutive nucleotide base pairs of the 3' terminal region beginning at

CC the 3' terminal base pair, and at least 125 consecutive nucleotide base

CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI

CC to the end of the piggyBac molecule. The DNA molecule in the transposon

CC piggyBac is useful for transferring genes into host cells or embryos for

CC transforming the cells of embryos. The transformed cells or embryos are

CC useful for developing or making transgenic organisms. This sequence

CC represents plasmid used in the creation of minimal sequence eukaryotic

CC transformation vector piggyBac

Sequence 4613 BP, 1064 A, 1183 C, 1273 G, 1093 T, 0 U, 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 4613;
Best Local Similarity 100.0%; Pred. No. 6e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGATCCCATGCGTCAATTTTACGAGACTATCTTTTACGGTTATCTAGCTGCATCAGG 60
DB 294 GGATCCCATGCGTCAATTTTACGAGACTATCTTTTACGGTTATCTAGCTGCATCAGG 353
QY 61 ATCATATGCTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGCA 120
DB 354 ATCATATGCTGGGCTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGCA 413
QY 121 TCGGGAGAGAGAACCCGCTTTTTCGGCAAGTTGAAGCGGATGGAAGAGATTG 180
DB 414 TCGGGAGAGAGAACCCGCTTTTTCGGCAAGTTGAAGCGGATGGAAGAGATTG 180
QY 181 CCGAGATGACTGCTGCTGATTCAGCTTGAGCGAAGACGCTTACATGATGATTC 240
DB 474 CCGAGATGACTGCTGCTGATTCAGCTTGAGCGAAGACGCTTACATGATGATTC 240
QY 241 GGGAGAGTGTGGCCATGACAGCCCTTTTAAAGCTGTTGCTTACAGCCAGCTGGATA 300
DB 534 GGGAGAGTGTGGCCATGACAGCCCTTTTAAAGCTGTTGCTTACAGCCAGCTGGATA 300
QY 301 CCAAGTTCGTGCGCGCTTTTTCGGACACAGTTCCGATGAGTGCAGCCCGAAGCGCATCAGCA 360
DB 594 CCAAGTTCGTGCGCGCTTTTTCGGACACAGTTCCGATGAGTGCAGCCCGAAGCGCATCAGCA 360
QY 361 ACCCGAACAATACCGCGCAGACCGGAACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

```

DB 654 ACCGGAATATACCGGCAACCGGAGTCCCGGCTGTGAGATTATGACAGC 713
 QY 421 GTGGGCGCTGGATATTATCTGACGAGACGGGTATCTGGTGTGATGCCAGAAAT 480
 DB 714 GTGGGCGCTGGATATTATCTGACGAGACGGGTATCTGGTGTGATGCCAGAAAT 773
 QY 481 GGACATGATATCCCGTGAATTAACCGGCGGCGGCTCTCTTCACTTACGTTTTGAA 540
 DB 774 GGACATGATATCCCGTGAATTAACCGGCGGCGGCTCTCTTCACTTACGTTTTGAA 833
 QY 541 CCGTGAAGAGAGGAG 600
 DB 834 CCGTGAAGAGAGGAG 893
 QY 601 AGATGCTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 894 AGATGCTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 953
 QY 661 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 707
 DB 954 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 1000

RESULT 3

ACAS5361/c
 ID ACAS5361 standard; DNA; 8999 BP.

ACAS5361;

06-JUN-2003 (first entry)

Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.

PiggyBac; transposon; eukaryotic transformation vector; ds;

transfected cell; transformed embryo; transgenic; plasmid; cyclic;

Synthetic.

US2002173634-A1.

21-NOV-2002.

30-OCT-2001; 2001US-00001189.

31-OCT-2000; 2000US-0244677P.

01-NOV-2000; 2000US-0244984P.

(FRASER M J.

(LIXX/) LI X.

(BEAM/) BEAM T.

(HUA-VAN A.

Fraser MJ, Li X, Beam T, Hua-Van A;

WPI; 2003-352597/33.

New DNA molecule in the transposon piggyBac, useful for transferring

genes into host cells or embryos for transforming the cells of embryos

that can be used in making transgenic organisms.

The invention describes a DNA molecule comprising at least 163

consecutive nucleotide base pairs of the 3' terminal region beginning at

the 3' terminal base pair, and at least 125 consecutive nucleotide base

pairs of the 5' terminal region beginning at the 5' terminal base pair of

the piggyBac molecule. The region extends from the restriction site SacI

to the end of the piggyBac molecule. The DNA molecule in the transposon

piggyBac is useful for transferring genes into host cells or embryos for

transforming the cells of embryos. The transformed cells or embryos are

useful for developing or making transgenic organisms. This sequence

CC represents plasmid used in the creation of minimal sequence eukaryotic

transformation vector piggyBac

Sequence 8999 BP; 2180 A; 2416 C; 2342 G; 2061 T; 0 U; 0 Other;

Query Match 99.8%; Score 705.4; DB 10; Length 8999;

Best Local Similarity 99.9%; Pred. No. 2.7e-212;

Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGATCCCATGCGTCAATTTTAAAGCAGACTATCTTTCTAGAGTTAATCTAGCTGATCAG 60
 DB 8950 GGATCCCATGCGTCAATTTTAAAGCAGACTATCTTTCTAGAGTTAATCTAGCTGATCAG 8891
 QY 61 ATCATATCTGCGGCTCTTTTCCGCTCACTATCCGCCCAAGTGGCGCTATCTGGGCA 120
 DB 8890 ATCATATCTGCGGCTCTTTTCCGCTCACTATCCGCCCAAGTGGCGCTATCTGGGCA 8831
 QY 121 TCGGGAGAGAAAG 180
 DB 8830 TCGGGAGAGAAAG 8771
 QY 181 CCGAGATGACTGCTGCTGATTTGACGTTGACGAAACGACGCTTACATGATGATTC 240
 DB 8770 CCGAGATGACTGCTGCTGATTTGACGTTGACGAAACGACGCTTACATGATGATTC 8711
 QY 241 GGGAGAGTGGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 DB 8710 GGGAGAGTGGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8651
 QY 301 CCAAGTTCGTCGCGGCTTTTCCGAGACAGAGTTCCGGATGTCAGCCGAGAGGATCAGCA 360
 DB 8650 CCAAGTTCGTCGCGGCTTTTCCGAGACAGAGTTCCGGATGTCAGCCGAGAGGATCAGCA 8591
 QY 361 ACCGGAACAATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 DB 8590 ACCGGAACAATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8531
 QY 421 GTGGGCGCTGGATATTATGCTCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 DB 8530 GTGGGCGCTGGATATTATGCTCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8471
 QY 481 GGACATGATATCCCGTGAATTAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 DB 8470 GGACATGATATCCCGTGAATTAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8411
 QY 541 CCGTGAAG 600
 DB 8410 CCGTGAAG 8351
 QY 601 AGATGCTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 8350 AGATGCTGACAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8291
 QY 661 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 707
 DB 8290 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 8244

RESULT 4

ACAS5362/c
 ID ACAS5362 standard; DNA; 9009 BP.

ACAS5362;

06-JUN-2003 (first entry)

Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.

PiggyBac; transposon; eukaryotic transformation vector; ds;

transfected cell; transformed embryo; transgenic; plasmid; cyclic;

Synthetic.

XX US2002173634-A1.
 XX 21-NOV-2002.
 XX 30-OCT-2001; 2001US-00001189.
 XX 31-OCT-2000; 2000US-0244677P.
 XX 01-NOV-2000; 2000US-0244984P.
 XX (FRAS/) FRASER M J.
 XX (LIXX/) LI X.
 XX (BEAM/) BEAM T.
 XX (HUA/V/) HUA-VAN A.
 XX PI
 XX Fraser MJ, Li X, Beam T, Hua-Van A;
 XX WPI; 2003-352597/33.
 XX
 XX New DNA molecule in the transposon piggyBac, useful for transferring
 XX genes into host cells or embryos for transforming the cells of embryos
 XX that can be used in making transgenic organisms.
 XX
 XX Example 11; Fig 13(B); 151pp; English.
 XX
 XX The invention describes a DNA molecule comprising at least 163
 XX consecutive nucleotide base pairs of the 3' terminal region beginning at
 XX the 3' terminal base pair, and at least 125 consecutive nucleotide base
 XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 XX the piggyBac molecule. The region extends from the restriction site SacI
 XX to the end of the piggyBac molecule. The DNA molecule in the transposon
 XX piggyBac is useful for transferring genes into host cells or embryos for
 XX transforming the cells of embryos. The transformed cells or embryos are
 XX useful for developing or making transgenic organisms. This sequence
 XX represents plasmid used in the creation of minimal sequence eukaryotic
 XX transformation vector piggyBac
 XX
 XX Sequence 9009 BP; 2179 A; 2429 C; 2341 G; 2058 T; 0 U; 2 Other;

Query Match 99.8%; Score 705.4; DB 10; Length 9009;
 Best Local Similarity 99.9%; Pred. No. 2,7e-212;
 Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGATCCCATGCGTCAATTTTACGACACTATCTTTCTAGGTTATCTAGCTGCATCAG 60
 DB 8960 GGATCCCATGCGTCAATTTTACGACACTATCTTTCTAGGTTATCTAGCTGCATCAG 60
 QY 61 ATCATATCGTCGGGCTCTTTTCCGGCTCACTCATCCCAAGCTGCGCTATCTGGGCA 120
 DB 8900 ATCATATCGTCGGGCTCTTTTCCGGCTCACTCATCCCAAGCTGCGCTATCTGGGCA 120
 QY 121 TCGGGGAGGAAAGAGCCCGTCCCTTTTCCCGAGGTTGAAGCGGATGGAAAGATTG 180
 DB 8840 TCGGGGAGGAAAGAGCCCGTCCCTTTTCCCGAGGTTGAAGCGGATGGAAAGATTG 180
 QY 181 CCGAGATGATGCTGCTGATGATGAGTGAAGGAAAGCAAGCTTTACATGATGATTC 240
 DB 8780 CCGAGATGATGCTGCTGATGATGAGTGAAGGAAAGCAAGCTTTACATGATGATTC 240
 QY 241 GGGAAAGTGTGGCCATGACGCTTTTAAACGTTGATGTTCTGTTAGGCAACCTGGGATA 300
 DB 8720 GGGAAAGTGTGGCCATGACGCTTTTAAACGTTGATGTTCTGTTAGGCAACCTGGGATA 300
 QY 301 CCAAGTGTGCGGCGCTTTTCCGAGACAGTTCCGATGCTGACCCGGAAGGCATCAGCA 360
 DB 8660 CCAAGTGTGCGGCGCTTTTCCGAGACAGTTCCGATGCTGACCCGGAAGGCATCAGCA 360
 QY 361 ACCCGAACAATACCGCGACAGCCGGAATGCGTGGCGGTGTGAAGTTAATGACGCG 420
 DB 8600 ACCCGAACAATACCGCGACAGCCGGAATGCGTGGCGGTGTGAAGTTAATGACGCG 420
 QY 421 GTGCGGCGTGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8541
 DB 421 GTGCGGCGTGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8541

DB 8540 GTGCGGCGCTGGGATATTACGTACAGAGAGACGGGTATCTGCTGATGCCGAGAAAT 8481
 QY 481 GGACATGGATACCCCGGAGTTACCCGCGCGCGCGCTCTGTTCAATCACTTTTGAAC 540
 DB 8480 GGACATGGATACCCCGGAGTTACCCGCGCGCGCGCTCTGTTCAATCACTTTTGAAC 540
 QY 541 CCGTGAGGACGCGGACACTCGCGGTGCAAAATGTTTTCACGCGGATGAGCAGATGA 600
 DB 8420 CCGTGAGGACGCGGACACTCGCGGTGCAAAATGTTTTCACGCGGATGAGCAGATGA 600
 QY 601 AGATGCTGACACGCTGCAAGAACACGACACTGATTAACCTTGAAGAATATCATATTG 660
 DB 8360 AGATGCTGACACGCTGCAAGAACACGACACTGATTAACCTTGAAGAATATCATATTG 660
 QY 661 TGACGTACGTTAAAGATTAATCATGCGTAAATTTGACGATGGGATCC 707
 DB 8300 TGACGTACGTTAAAGATTAATCATGCGTAAATTTGACGATGGGATCC 8254

RESULT 5
 ACAS5363/c
 ID ACAS5363 standard; DNA; 9012 BP.
 XX ACAS5363;
 AC
 XX 06-JUN-2003 (first entry)
 XX
 XX Transformation vector piggyBAC related plasmid p(pZ)-Bac-Egfp.
 DE
 XX
 XX piggyBac; transposon; eukaryotic transformation vector; ds;
 KM transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 KW circular.
 XX
 XX Synthetic.
 OS
 XX
 XX US2002173634-A1.
 XX
 XX 21-NOV-2002.
 XX
 XX 30-OCT-2001; 2001US-00001189.
 XX
 XX 31-OCT-2000; 2000US-0244677P.
 XX 01-NOV-2000; 2000US-0244984P.
 XX
 XX (FRAS/) FRASER M J.
 XX (LIXX/) LI X.
 XX (BEAM/) BEAM T.
 XX (HUA/V/) HUA-VAN A.
 XX
 XX Fraser MJ, Li X, Beam T, Hua-Van A;
 XX WPI; 2003-352597/33.
 XX
 XX New DNA molecule in the transposon piggyBac, useful for transferring
 XX genes into host cells or embryos for transforming the cells of embryos
 XX that can be used in making transgenic organisms.
 XX
 XX Example 11; Fig 14(B); 151pp; English.
 XX
 XX The invention describes a DNA molecule comprising at least 163
 XX consecutive nucleotide base pairs of the 3' terminal region beginning at
 XX the 3' terminal base pair, and at least 125 consecutive nucleotide base
 XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 XX the piggyBac molecule. The region extends from the restriction site SacI
 XX to the end of the piggyBac molecule. The DNA molecule in the transposon
 XX piggyBac is useful for transferring genes into host cells or embryos for
 XX transforming the cells of embryos. The transformed cells or embryos are
 XX useful for developing or making transgenic organisms. This sequence
 XX represents plasmid used in the creation of minimal sequence eukaryotic
 XX transformation vector piggyBac
 XX
 XX Sequence 9012 BP; 2180 A; 2419 C; 2353 G; 2060 T; 0 U; 0 Other;

QY 541 CCGTGGAGGACGGGCGAGCTCCGGGTGCAAAATGTGTTTACAGCGGTGATGGACGATGA 600
DB 1253 CCGTGGAGGACGGGCGAGCTCCGGGTGCAAAATGTGTTTACAGCGGTGATGGACGATGA 1312
QY 601 AGATGCTGACACGCTGACAGACGACGCTAGATTACCTTAAGAAAGATATATATTTG 660
DB 1313 AGATGCTGACACGCTGACAGACGACGCTAGATTACCTTAAGAAAGATATATATTTG 1372
QY 661 TGACGTACGTTAAAGATATATCATGCGTAAATTTGACGATGGATCC 707
DB 1373 TGACGTACGTTAAAGATATATCATGCGTAAATTTGACGATGGATCC 1419

RESULT 7
ID ACA55367 standard; DNA; 4943 BP.
AC ACA55367;
XX
XX 06-JUN-2003 (first entry)
XX
XX
XX Transformation vector piggyBAC related plasmid PBS-ITR-EGFP.
XX PiggyBac; transposon; eukaryotic transformation vector; ds;
XX transformed cell; transformed embryo; transgenic; plasmid; cyclic;
OS Synthetic.
XX US2002173634-A1.
XX
XX 21-NOV-2002.
XX
XX 30-OCT-2001; 2001US-00001189.
XX
XX 31-OCT-2000; 2000US-0244677P.
XX 01-NOV-2000; 2000US-0244984P.
XX
XX (FRAS/) FRASER M J.
XX (LIXX/) Li X.
XX (BEAM/) BEAM T.
XX (HUAV/) HUA-VAN A.
XX
XX Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
XX genes into host cells or embryos for transforming the cells of embryos
XX that can be used in making transgenic organisms.
PS Example 6; Fig 18 (B); 151pp; English.

CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC the piggyBAC molecule. The region extends from the restriction site SacI
CC to the end of the piggyBAC molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
XX
SQ Sequence 4943 BP; 1221 A; 1305 C; 1245 G; 1172 T; 0 U; 0 Other;

Query Match 96.9%; Score 685; DB 10; Length 4943;
Best Local Similarity 99.2%; Pred. No. 5.9e-206;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 GATCCCATGCGTCAATTTTACGACGACTATCTTCTAGGTTAATCATGCTGATCAGG 60

DB 719 GGATCCCATGCGTCAATTTTACGACGACTATCTTCTAGGTTAATCATGCTGATCAGG 778
QY 61 ATCATATGTTGGGTTCTTTTCCGGCTCAGTCATCGCCAGACTGCTATCTGGGCA 120
DB 779 ATCATATGTTGGGTTCTTTTCCGGCTCAGTCATCGCCAGACTGCTATCTGGGCA 838
QY 121 TCGGGAGGAAGAACCCGCTTTTCCGGCAGAGGTGAAGGGCATGGAAGAGTTTG 180
DB 839 TCGGGAGGAAGAACCCGCTTTTCCGGCAGAGGTGAAGGGCATGGAAGAGTTTG 898
QY 181 CCGAGATGACCTGCTGCTGATTTGACGTTAGCGAAAGCAAGTTTACATGATGATTC 240
DB 899 CCGAGATGACCTGCTGCTGATTTGACGTTAGCGAAAGCAAGTTTACATGATGATTC 958
QY 241 GGGAGGTTGTGGCAGTCAGACGCTTTTAAAGGTGAAGCTGTTTACGACCTGGGATTA 300
DB 959 GGGAGGTTGTGGCAGTCAGACGCTTTTAAAGGTGAAGCTGTTTACGACCTGGGATTA 1018
QY 301 CCAAGTTCTGCGGGCTTTTCCGAGACACAGTTCCGATGATGAGCCGAGAGCATCAGCA 360
DB 1019 CCAAGTTCTGCGGGCTTTTCCGAGACACAGTTCCGATGATGAGCCGAGAGCATCAGCA 1078
QY 361 ACCCGAACAATACCGGCGACACCGGAACTGCGCGGCTGGTGCAGTTTATGACAGCG 420
DB 1079 ACCCGAACAATACCGGCGACACCGGAACTGCGCGGCTGGTGCAGTTTATGACAGCG 1138
QY 421 GTGCGGCGCTGGGATTTTACGTCAGCGAGACGCGGTATCTGCTGATGCCGAGAAAT 480
DB 1139 GTGCGGCGCTGGGATTTTACGTCAGCGAGACGCGGTATCTGCTGATGCCGAGAAAT 1198
QY 481 GGACATGTATACCCCGTAGTTACCCGCGGCGCGCTGTCATTCAAGTTTGAAC 540
DB 1199 GGACATGTATACCCCGTAGTTACCCGCGGCGCGCTGTCATTCAAGTTTGAAC 1252
QY 541 CCGTGGAGGACGGGCGAGCTCCGGTCAAAATGTGTTTACAGCGATGGACGATGA 600
DB 1253 CCGTGGAGGACGGGCGAGCTCCGGTCAAAATGTGTTTACAGCGATGGACGATGA 1312
QY 601 AGATGCTGACACGCTGACAGACGACGCTAGATTACCTTAAGAAAGATATATTTG 660
DB 1313 AGATGCTGACACGCTGACAGACGACGCTAGATTACCTTAAGAAAGATATATTTG 1372
QY 661 TGACGTACGTTAAAGATATATCATGCGTAAATTTGACGATGGATCC 707
DB 1373 TGACGTACGTTAAAGATATATCATGCGTAAATTTGACGATGGATCC 1419

RESULT 8
ID ACA55368 standard; DNA; 4944 BP.
AC ACA55368;
XX
XX 06-JUN-2003 (first entry)
XX
XX
XX Transformation vector piggyBAC related plasmid PBS-ITR-EGFP.
XX PiggyBac; transposon; eukaryotic transformation vector; ds;
XX transformed cell; transformed embryo; transgenic; plasmid; cyclic;
OS Synthetic.
XX
XX
XX US2002173634-A1.
XX
XX 21-NOV-2002.
XX
XX 30-OCT-2001; 2001US-00001189.
XX
XX 31-OCT-2000; 2000US-0244677P.
XX 01-NOV-2000; 2000US-0244984P.
XX

PT	Prepared from various human tissues, for diagnosis, treatment of cancer,
PT	neurological, inflammatory disorders and for use in arrays for detection
XX	Claim 1, SEQ ID NO 158, 153bp, English.
CC	Sequences AAS4576-AAS44919 represent full-length polynucleotides and
CC	contig polynucleotides encoding polypeptides of the invention. The DNA
CC	and protein sequences are useful for the treatment, diagnosis and
CC	prevention of various types of disorder in a mammalian subject such as a
CC	human, dog, monkey, mouse, hamster or rat. The disorders include such as a
CC	such as leukemias, lymphoma and neuroblastoma, autoimmune disorders such as
CC	diabetes mellitus, allergic rhinitis, asthma and eczema, rheumatoid arthritis,
CC	disorders such as Parkinson's disease, Alzheimer's disease, nervous system
CC	chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC	Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC	disease, ischemia-reperfusion injury, shock, sepsis and inflammatory
CC	bowel disease. The sequences exhibit activity relating to angiogenesis,
CC	cell proliferation, cell differentiation, stem cell growth factor,
CC	activin or inhibin. Therefore, they can be used to manipulate stem cells
CC	in culture to give rise to neuroepithelial cells that can be used to
CC	augment or replace cells damaged by illness, accidental damage or genetic
CC	disorders. The sequences may also be used for regeneration of bone.
CC	Note: Some sequences for this patent did not form part of the printed
CC	specification, but were obtained in electronic format directly from WIPO
CC	at ftp://www.int/pub/published_pat_sequences
XX	
XX	Sequence 4164 BP; 906 A; 1123 C; 1307 G; 828 T; 0 U; 0 Other;
XX	
XX	Query Match 82.1%; Score 580.6; DB 4; Length 4164;
XX	Best Local Similarity 99.3%; Pred. No. 6.9e-173;
XX	Matches 583; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
YY	
YY	45 ATCTAGCTGATCAAGATATCATATTCGTGGGCTCTTTTCCGGCTAGCATGCGCCAAAC 104
YY	178 ATCCAGCGCATCAGATCATATTCGTGGGCTCTTTTCCGGCTAGCATGCGCCAAAC 237
YY	105 TGGGCGCTATCTGGGCTATCGGGAGAGAAAGCCGCTTCTTCCCGAGAGTTGAACG 164
YY	238 TGGGCGCTATCTGGGCTATCGGGAGAGAAAGCCGCTTCTTCCCGAGAGTTGAACG 164
YY	165 GCATGGAAGAAGTTTCCGAGAGATCATCTGCTGTGCATTTAGCTTGAGCAAAAGCAGC 297
YY	298 GCATGGAAGAAGTTTCCGAGAGATCATCTGCTGTGCATTTAGCTTGAGCAAAAGCAGC 224
YY	225 TTACACATATATATTCGGGAAGGTGTGGCATGACGCCCTTAAACGCTGTAACCTGTTCT 284
YY	358 TTTCACATATATATTCGGGAAGGTGTGGCATGACGCCCTTAAACGCTGTAACCTGTTCT 417
YY	285 CAGGCGCACTGGGATACCAAGTTGCTGGCGGCTTTTCCGAGCAACAGTTCCGATGCTACG 344
YY	418 CAGGCGCACTGGGATACCAAGTTGCTGGCGGCTTTTCCGAGCAACAGTTCCGATGCTACG 344
YY	345 CCGAAGCGCATACACACACCCGAAACAATATCCGCGCAGCGCGGAATGCGCGGCGGCTGTG 404
YY	478 CCGAAGCGCATACACACACCCGAAACAATATCCGCGCAGCGCGGAATGCGCGGCGGCTGTG 404
YY	405 CAGATTATATACACCGGTGCGCGCTGGGATATTCGTCACGAGAGACGGGTATCTCTGAC 464
YY	538 CAGATTATATACACCGGTGCGCGCTGGGATATTCGTCACGAGAGACGGGTATCTCTGAC 464
YY	465 TGAGTCCGAGAAATGAGCATGATATCCCGTGAATTTACCGGCGGCGCGCTGCTTC 597
YY	598 TGAGTCCGAGAAATGAGCATGATATCCCGTGAATTTACCGGCGGCGCGCTGCTTC 524
YY	525 ATTACAGCTTTTGAACCCGTGAGAGACGGGCAAGCTGCGGCTGCAAAATGTGTTTAAACG 657
YY	658 ATTACAGCTTTTGAACCCGTGAGAGACGGGCAAGCTGCGGCTGCAAAATGTGTTTAAACG 584
YY	585 GTGATGAGCAGATGAGATGCTTGACACCTTGACAAACAGCAGCT 631

RESULT 11

AB878924/c

ID AB878924 standard; DNA; 26565 BP.

AC AB878924;

DT 17-DEC-2002 (first entry)

DE E. coli CFT073 genomic sequence #91.

XX Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KM urinary tract infection; open reading frame; ORF; uropathogenic;
KM antibacterial; atrophic; nephrotoxic; gene; ds.

OS Escherichia coli.

XX MO300259320-A2.

XX 01-AUG-2002.

XX 19-OCT-2001; 2001WO-US046833.

XX 19-OCT-2000; 2000US-0242412P.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

XX Blattner FR, Welch RA, Burland VD;

XX WPI; 2002-691532/74.

PT New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT useful for preventing or treating E. coli CFT073 infection in humans or
PT livestock.

PS Claim 1; Page 276-291; 765dp; English.

XX The present invention relates to polynucleotide sequences from the genome
CC of the pathogenic Escherichia coli strain CFT073. Almost all the
CC sequences present in E. coli CFT073 are absent in the previously
CC sequenced laboratory strain K-12. The polynucleotide sequences of the
CC invention are useful for preventing, diagnosing or treating E. coli
CC CFT073 infection in humans or livestock. The polynucleotide sequences are
CC useful for preventing urinary tract infections and pyelonephritis.
CC Likewise, the polypeptides encoded by the different open reading frames
CC (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC coli strains. AB878834-AB879085 represent genomic sequences from E. coli
CC strain CFT073

SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

Query Match 80.1%; Score 566.2; DB 6; Length 26565;

Best Local Similarity 97.8%; Pred. No. 6.5e-168;

Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCAGATCATATGTCGGGCTTTTTCGCGCTCAGTCATGCCCAAGC 104

DB 3221 ATTCAGCTGATCAGATCATATGTCGGGCTTTTTCGCGCTCAGTCATGCCCAAGC 3162

QY 105 TGGCGCTATCGGCGATCGGGAGAGAAAGCCGTCCTTTTCCCGAGAGTTGAAGCG 164

DB 3161 TGGCGCTATCGGCGATCGGGAGAGAAAGCCGTCCTTTTCCCGAGAGTTGAAGCG 3102

QY 165 GCATGGAAGAGTTTGGCGAGATGACTGCTGCTGATTTGACGTTGACGAAAGCAGC 224

DB 3101 GCATGGAAGAGTTTGGCGAGAGATGACTGCTGCTGATTTGACGTTGACGAAAGCAGC 3042

QY 225 TTATACCTATGATTTGGGAAAGTTGGCCATGACAGCCCTTTAAAGCTGTAAGCTTTGTT 284

DB 3041 TTATACCTATGATTTGGGAAAGTTGGCCATGATGCTTTAAAGCTGTAAGCTTTGTT 2982

QY 285 CAGGCACTGGATACCAAGTTGTCGGGGCTTTTCCGGAACAAGTCCGAGTGTGACG 344

DB 2981 CAGGCACTGGATACCAAGTTGTCGGGGCTTTTCCGGAACAAGTCCGAGTGTGACG 2922

QY 345 CCGAAGCATACGAAACCCGAAACATACCGGCGACAGCCGGAACTCCCGTGGCGGTGTG 404

DB 2921 CCGAAGCATACGAAACCCGAAACATACCGGCGACAGCCGGAACTCCCGTGGCGGTGTG 2862

QY 405 CAGATTATGACAGCGGTGGCGGCTGGGATTTACGTCAGCGAGACGGGTATCCGTGGC 464

DB 2861 CAGATTATGACAGCGGTGGCGGCTGGGATTTACGTCAGCGAGACGGGTATCCGTGGC 2802

QY 465 TGGATGCCGAGAAATGACATGATATCCCGTGAATTACCCGGGAGCGGCTCGTTG 524

DB 2801 TGGATGCCGAGAAATGACATGATATCCCGTGAATTACCCGGGAGCGGCTCGTTG 2742

QY 525 ATTACGTTTTTGAACCCGTGAGAGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 584

DB 2741 ATTACGTTTTTGAACCCGTGAGAGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 2682

QY 585 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 631

DB 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

QY 2681 GTGATGAGAGATGAAGATGCTTGACACGCTGAGAAACAGCAGCT 2635

```
Db 3221 ATCAGCTGATACGATCATATCGTGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 3162
Qy 105 TGGCGCTATCTGGGCATCGGAGGAGAAAGCCCGTCTTTCCCGGAGTTGAACG 164
Db 3161 TGGCGCTATCTGGGCATCGGAGGAGAAAGCCCGTCTTTCCCGGAGTTGAACG 3102
Qy 165 GCATGGAAGAAGTTGGCCGAGATACCTGCTGTCATGACGTTGAGCCGAAAACGACG 224
Db 3101 GCATGGAAGAAGTTGGCCGAGATACCTGCTGTCATGACGTTGAGCCGAAAACGACG 3042
Qy 225 TTACCATGATGATTTGGGAAAGTGTGGCATGACGCTTTAAACGGTGAATCTTCTT 284
Db 3041 TTCACCATGATGATTTGGGAAAGTGTGGCATGACGCTTTAAACGGTGAATCTTCTT 2982
Qy 285 CAGGCCACCTGGGATACCAAGTCTCCGCTTTTCCGGAACAGATTCGCGATGTGACG 344
Db 2981 CAGGCCACCTGGGATACCAAGTCTCCGCTTTTCCGGAACAGATTCGCGATGTGACG 2922
Qy 345 CCGAAGCGCATCAGCAACCCGAAACATATACCGGACGACGCGGAACTGCGCTGCTG 404
Db 2921 CCGAAGCGCATCAGCAACCCGAAACATATACCGGACGACGCGGAACTGCGCTGCTG 2862
Qy 405 CAGATTATGACAGCGGCTGGGGGCTGGCATATTACGTGACGAGGACGGATCTCTGAC 464
Db 2861 CAGATTATGACAGCGGCTGGGGGCTGGCATATTACGTGACGAGGACGGATCTCTGAC 2802
Qy 465 TGGATGCGCGAAGATGACATGATATACCCGTAATACCGGGGCGGCGCTCGTTC 524
Db 2801 TGGATGCGCGAAGATGACATGATATACCCGTAATACCGGGGCGGCGCTCGTTC 2742
Qy 525 ATTCAGGTTTGAACCCGTTGAGAGACGGGACGACTCGCGTGCATAATGTGTTTACAGC 584
Db 2741 ATTCAGGTTTGAACCCGTTGAGAGACGGGACGACTCGCGTGCATAATGTGTTTACAGC 2682
Qy 585 GTGATGAGACAGATGAAAGTGTCTGACACGCTGACAGAACGACGAGCT 631
Db 2681 GTGATGAGACAGATGAAAGTGTCTGACACGCTGACAGAACGACGAGCT 2635

RESULT 13
ACD19059
ID ACD19059 standard; DNA; 46819 BP.
XX
AC ACD19059;
XX
DT 27-OCT-2003 (revised)
DT 21-AUG-2003 (first entry)
XX
DE E. coli 0157 unique DNA sequence OZID_72.
XX
KW OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
XX food poisoning.
XX
OS Escherichia coli; strain 0157:H7.
XX
PN US2003023075-A1.
XX
PD 30-JAN-2003.
XX
PF 01-APR-2002; 2002US-00114170.
XX
PR 04-DEC-1998; 98US-0110955P.
XX
PR 03-DEC-1999; 99US-00453702.
XX
PA (BLAT/) BLATTNER F R.
PA (BURL/) BURLAND V D.
PA (PERN/) PERNA N T.
PA (PLUNK/) PLUNKETT G.
PA (WELC/) WELCH R.
XX
PI Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;
```

```
DR WPI; 2003-479497/45.
XX
PT New DNA sequences from Escherichia coli strain 0157:H7, useful for
PT detecting E. coli 0157:H7 in a sample, or in designing diagnostic probes
PT which can be used to distinguish strain 0157:H7 from strain K12 using
PT molecular techniques.
XX
PS Claim 16; SEQ ID NO 72; 33pp; English.
XX
CC The invention relates to an isolated DNA molecule comprising an E. coli
CC strain 0157:H7 sequence selected from a clonal library of a locus of
CC a urease gene cluster, a RTX toxin-like gene cluster, a locus of
CC enterocyte effacement and 2 genes from its associated lymphocytic phage
CC 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC 0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC (which can develop into haemolytic uraemic syndrome). Also included are
CC an isolated DNA molecule comprising a nucleotide sequence identical to at
CC least 25 contiguous nucleotides contained in DNA sequences selected from
CC ACD18988-ACD19242 (being 25 E. coli 0157 DNA sequences which are not
CC found in E. coli K12), a recombinant DNA construction comprising the DNA
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPO at
CC segdata.uspto.gov/sequence.html?docid=2003023075 (Updated on 27-OCT-2003
XX to standardise OS field)
XX
SQ Sequence 46819 BP; 11858 A; 10855 C; 12774 G; 11317 T; 0 U; 15 Other;
```

```
Query Match
Beet Local Similarity 80.1%; Score 566.2; DB 9; Length 46819;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
Qy 45 ATCTAGCTGATCAGATCATATCGTGGGCTCTTTTCCGCTCAGTCATCGCCCAAGC 104
Db 14795 ATCCAGCTGATCAGATCATATCGTGGGCTCTTTTCCGCTCAGTCATCGCCCAAGC 14854
Qy 105 TGGCGCTATCTGGGCATCGGAGGAGAAAGCCCGTCTTTCCCGGAGTTGAACG 164
Db 14855 TGGCGCTATCTGGGCATCGGAGGAGAAAGCCCGTCTTTCCCGGAGTTGAACG 14914
Qy 165 GCATGGAAGAAGTTGGCCGAGATACCTGCTGTCATGACGTTGAGCCGAAAACGACG 224
Db 14915 GCATGGAAGAAGTTGGCCGAGATACCTGCTGTCATGACGTTGAGCCGAAAACGACG 14974
Qy 225 TTTCACATATGATTTGGGAAAGTGTGACATGACGCTTTAAACGGTGAATCTTCTT 284
Db 14975 TTTCACATATGATTTGGGAAAGTGTGACATGACGCTTTAAACGGTGAATCTTCTT 15034
Qy 285 CAGGCCACCTGGGATACCAAGTCTCCGCTTTTCCGGAACAGATTCGCGATGTGACG 344
Db 15035 CAGGCCACCTGGGATACCAAGTCTCCGCTTTTCCGGAACAGATTCGCGATGTGACG 15094
Qy 345 CCGAAGCGCATCAGCAACCCGAAACATATACCGGACGACGCGGAACTGCGCTGCTG 404
Db 15095 CCGAAGCGCATCAGCAACCCGAAACATATACCGGACGACGCGGAACTGCGCTGCTG 15154
Qy 405 CAGATTATGACAGCGGCTGGGGGCTGGCATATTACGTGACGAGGACGGATCTCTGAC 464
Db 15155 CAGATTATGACAGCGGCTGGGGGCTGGCATATTACGTGACGAGGACGGATCTCTGAC 15214
Qy 465 TGGATGCGCGAAGATGACATGATATACCCGTAATACCGGGGCGGCGCTCGTTC 524
Db 15215 TGGATGCGCGAAGATGACATGATATACCCGTAATACCGGGGCGGCGCTCGTTC 15274
Qy 525 ATTCAGGTTTGAACCCGTTGAGAGACGGGACGACTCGCGTGCATAATGTGTTTACAGC 584
```

Db 15275 ATTCAGCTCTTGAACCCGCGAGAGCGGGAGACCCCGCGGTGCAATGTGTTTACAGC 15334
Qy 585 GTGATGGAGCAGATGAAGATGCTGACACGCTGCAGAAACAGCAGCT 631
Db 15335 GTGATGGAGCAGATGAAGATGCTGACACGCTGCAGAAACAGCAGCT 15381

RESULT 14
ID ADC00585 standard; DNA; 46897 BP.
AC ADC00585;
DT 04-DEC-2003 (first entry)
DE Enterohaemorrhagic E. coli O157:H7-specific nucleic acid SEQ ID NO: 630.
KM ds; gene; enterohaemorrhagic; anti-bacterial.
XX Escherichia coli; O157:H7.
OS
PN JF2002355074-A.
PD 10-DEC-2002.
PE 24-JAN-2002; 2002JF-00015959.
PR 24-JAN-2001; 2001JF-00112010.
XX
XX (UYTS-) UNIV TSUKUBA.
DR WPI; 2003-451640/43.
XX
PT Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
XX and a polypeptide and its use, a polypeptide, a vector and a host cell.
PS Claim 2; SEQ ID NO 630; 2067pp; Japanese.
XX
CC The invention relates to a novel enterohaemorrhagic Escherichia coli
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present
CC sequence represents an E. coli O157:H7-specific nucleic acid of the
CC invention.
XX
SQ Sequence 46897 BP; 11872 A; 10948 C; 12731 G; 11346 T; 0 U; 0 Other;

Query Match 79.9%; Score 564.6; DB 10; Length 46897;
Best Local Similarity 97.6%; Pred. No. 2.8e-167;
Matches 573; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGATCATATGTCGCGGCTCTTTTCCGCGTCAGTCATCGCCCAAGC 104
Db 14795 ATCCAGCTGCATCAGATCATATGTCGCGGCTCTTTTCCGCGTCAGTCATCGCCCAAGC 14854
Qy 105 TGGGCGCTATCTGGGCGATCGGGAGAGAAAGCCCGTGCCTTTTCCCGGAGTTGAAGCG 164
Db 14855 TGGGCGCTATCTGGGCGATCGGGAGAGAAAGCCCGTGCCTTTTCCCGGAGTTGAAGCG 14914
Qy 165 GCATGGAAGAGTTTGGCGAGATGACTGCTGCATTGAGCTTGAGCGAAACGACG 224
Db 14915 GCATGGAAGAGTTTGGCGAGATGACTGCTGCATTGAGCTTGAGCGAAACGACG 14974
Qy 225 TTATACATGATGATTCGGGAAGGTGTGCCATGACCGCTTTAAACGTGAATGTTGCTT 284
Db 14975 TTATACATGATGATTCGGGAAGGTGTGCCATGACCGCTTTAAACGTGAATGTTGCTT 15034
Qy 285 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCCGAGACAGCTTCGATGCTCAGC 344
Db 15035 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCCGAGACAGCTTCGATGCTCAGC 15094
Qy 345 CCGAAGCGCATCAGCAACCCGAAACATATCCGGCGACGCGGAACTGCGCGGTGTG 404

Db 15095 CCGAAGCGCATCAGCAACCCGAAACATATCCAGCCACGCGGAACTGCGCGGTGTG 15154
Qy 405 CAGATTAAATGACACCGGTGCGCGCTGGGATATTACGTACAGCAGACGGGTATCTGGC 464
Db 15155 CAGATTAAATGACACCGGTGCGCGCTGGGATATTACGTACAGCAGACGGGTATCTGGC 15214
Qy 465 TGGATGCCCGCAGAAATGACATGATATACCCGTAGTTATCCCGCGCGCGCTCTGCTC 524
Db 15215 TGGATGCCCGCAGAAATGACATGATATACCCGTAGTTATCCCGCGCGCTCTGCTC 15274
Qy 525 ATTCACGTTTTTGAACCCGTGAGAGACGGGACACTCGCGGTGCAAAATGTTTACAGC 584
Db 15275 ATTCACGCTTTGAACCCGTGAGAGACGGGACACTCGCGGTGCAAAATGTTTACAGC 15334
Qy 585 GTGATGGAGCAGATGAAGATGCTGACACGCTGCAGAAACAGCAGCT 631
Db 15335 GTGATGGAGCAGATGAAGATGCTGACACGCTGCAGAAACAGCAGCT 15381

RESULT 15
ID ACD19238/c standard; DNA; 22306 BP.
AC ACD19238;
AC ACD19238;
DT 27-OCT-2003 (revised)
DT 21-AUG-2003 (first entry)
DE E. coli O157 unique DNA sequence OZID_251.
XX
XX OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KM Food poisoning.
XX
XX Escherichia coli; strain O157:H7.
OS
PN US2003023075-A1.
PD 30-JAN-2003.
XX
XX 01-APR-2002; 2002US-00114170.
PE 04-DEC-1998; 98US-0110955P.
PR 03-DEC-1999; 99US-00453702.
XX
XX (BLAT/) BLATTNER F R.
PA (BURL/) BURLAND V D.
PA (PERN/) PERNA N T.
PA (PLUN/) PLUNKETT G.
PA (WELC/) WELCH R.
PI Blatter FR, Burland VD, Perna NT, Plunkett G, Welch R;
XX
XX WPI; 2003-479497/45.
DR
XX
XX New DNA sequences from Escherichia coli strain O157:H7, useful for
PT detecting E. coli O157:H7 in a sample, or in designing diagnostic probes
PT which can be used to distinguish strain O157:H7 from strain K12 using
PT molecular techniques.
XX
PS Claim 16; SEQ ID NO 251; 33pp; English.

The invention relates to an isolated DNA molecule comprising an E. coli strain O157:H7 sequence selected from a clostridial cytotoxin-like gene, a urease gene cluster, a RTX toxin-like gene cluster, a locus of enterocyte effacement and 2 genes from its associated lymphocytic phage 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli O157:H7 can cause food poisoning, specifically acute haemorrhagic colitis (which can develop into haemolytic uraemic syndrome). Also included are an isolated DNA molecule comprising a nucleotide sequence identical to at least 25 contiguous nucleotides contained in DNA sequences selected from ACD18988-ACD19242 (being 25 E. coli O157 DNA sequences which are not found in E. coli K12), a recombinant DNA construction comprising the DNA

CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed O2ID 1-O2ID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?docid=2003023075 (updated on 27-Oct-2003
CC to standardise OS field)

XX
SQ Sequence 22306 BP; 4744 A; 6842 C; 5512 G; 5192 T; 0 U; 16 Other;

Query Match 79.4%; Score 561.4; DB 9; Length 22306;
Best Local Similarity 97.3%; Pred. No. 2e-166;

Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
QY 45 ATCTAGCTCATCAGATCATATCGCGGCTCTTTTCCGGCTCAGTCATGCCCCAAGC 104
DB 17193 ATCCAGTTGCATCAGATCATATCGCGGCTCTTTTCCGGCTCAGTCATGCCCCAAGC 17134
QY 105 TGGCGGTATCTGGGCAATCGGAGAGAAAGCCCGTCCCTTTCCCGAGGTTGAAGC 164
DB 17133 TGCGCTATCTGGGCAATCGGAGAGAAAGCCCGTCCCTTTCCCGAGGTTGAAGC 17074
QY 165 GCATGGAAGAAGTTGGCGAGAGATGCTGCTGCTGCTTTTAAAGGTTGAAGCAAGC 224
DB 17073 GCATGGAAGAAGTTGGCGAGATGCTGCTGCTGCTTTTAAAGGTTGAAGCAAGC 17014
QY 225 TTACCATGATGATTTGGGAAAGTGTGCGCATGCAAGCCCTTTAAAGGTTGAAGTGTGCTT 284
DB 17013 TTACCATGATGATTTGGGAAAGTGTGCGCATGCAAGCCCTTTAAAGGTTGAAGTGTGCTT 16954
QY 285 CAGGCCACTGGGATATCAATTCGCGGCTTTTCCGACACAGTTCGGATGTCAGC 344
DB 16953 CAGGCCACTGGGATATCCGCTCCCTCGGACTGTTCCGACACAGTTCGGATGTCAGC 16894
QY 345 CCGAAGCGCATCAGCAACCCGACATACCGGCGACAGCCGAACTGCCGCGGTTG 404
DB 16893 CCGAAGCGCATCAGCAACCCGACATACCGGCGACAGCCGAACTGCCGCGGTTG 16834
QY 405 CAGATTATGACAGCGGTCGCGCTGGGATTTAGCTCAGCGAGACGGGTAATCTGCGC 464
DB 16833 CAGATTATGACAGCGGTCGCGCTGGGATTTAGCTCAGCGAGACGGGTAATCTGCGC 16774
QY 465 TGGATGCCGAGAAATGACATGATACCCGCTGATTAACCGGCGGCGCGCTGCTTC 524
DB 16773 TGGATGCCGAGAAATGACATGATACCCGCTGATTAACCGGCGGCGCGCTGCTTC 16714
QY 525 ATTACGTTTTTGAACCCGTGAGAGACGGGCAACTCGCGGTGCAATGTGTTTACAGC 584
DB 16713 ATTACGTTTTTGAACCCGTGAGAGACGGGCAACTCGCGGTGCAATGTGTTTACAGC 16654
QY 585 GTGATGAGACAGATGAAGATGCTCGACAGCTGACAGACCGAGCT 631
DB 16653 GTGATGAGACAGATGAAGATGCTCGACAGCTGACAGACCGAGCT 16607
```

Search completed: January 26, 2005, 03:14:34
Job time : 338.971 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 08:26:03 ; Search time 378.34 Seconds
(without alignments)
10737.291 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707
Sequence: 1 ggaatccatgcgcgaatttc.....aaatgcacgacggatcc 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 287294193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	US-10-001-189-40	Sequence 40, App1
2	707	100.0	4613	US-10-001-189-46	Sequence 46, App1
3	707	100.0	8999	US-10-001-189-48	Sequence 48, App1
4	707	100.0	9012	US-10-001-189-49	Sequence 49, App1
5	707	100.0	9013	US-10-001-189-50	Sequence 50, App1
6	685	96.9	4943	US-10-001-189-54	Sequence 54, App1
7	685	96.9	4944	US-10-001-189-55	Sequence 55, App1
8	685	96.9	4944	US-10-001-189-56	Sequence 56, App1
9	566.2	80.1	26565	US-10-085-959-91	Sequence 91, App1
10	566.2	80.1	46819	US-10-114-170-72	Sequence 72, App1
11	561.4	79.4	22306	US-10-114-170-251	Sequence 251, App1
12	519.2	73.4	3662	US-10-001-189-41	Sequence 41, App1

C 13	511	72.3	4941	13	US-10-001-189-53	Sequence 53, App1
C 14	511	72.3	4951	13	US-10-001-189-51	Sequence 51, App1
C 15	511	72.3	4952	13	US-10-001-189-52	Sequence 52, App1
C 16	425.8	60.2	1427	18	US-10-363-345A-37517	Sequence 37517, A
C 17	425.8	60.2	1427	18	US-10-363-345A-37518	Sequence 37518, A
C 18	382.6	54.1	1427	18	US-10-363-345A-37519	Sequence 37519, A
C 19	382.6	54.1	1427	18	US-10-363-345A-37520	Sequence 37520, A
C 20	223	31.5	34063	14	US-10-114-170-96	Sequence 96, App1
C 21	222.2	31.4	72480	16	US-10-418-837-2	Sequence 2, App1
C 22	96	13.6	847	10	US-09-798-889-27	Sequence 27, App1
C 23	96	13.6	847	16	US-10-633-680-27	Sequence 27, App1
C 24	78	11.0	9984	13	US-10-001-189-63	Sequence 63, App1
C 25	72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, App1
C 26	40	5.7	43	13	US-10-001-189-19	Sequence 19, App1
C 27	40	5.7	3089	18	US-10-425-115-49726	Sequence 49726, A
C 28	39.8	5.6	676	16	US-10-425-115-91844	Sequence 91844, A
C 29	39.8	5.6	1673	16	US-10-425-114-21174	Sequence 21174, A
C 30	39.8	5.6	1796	16	US-10-425-114-29053	Sequence 29053, A
C 31	39.8	5.6	1872	18	US-10-425-115-91842	Sequence 91842, A
C 32	39	5.5	39	13	US-10-001-189-16	Sequence 16, App1
C 33	38.4	5.4	1473	16	US-10-282-122A-20134	Sequence 20134, A
C 34	36.8	5.2	4187	17	US-10-437-963-91699	Sequence 91699, A
C 35	36.2	5.1	1403	17	US-10-767-701-13724	Sequence 13724, A
C 36	36	5.1	1246	15	US-10-369-493-27114	Sequence 27114, A
C 37	35.4	5.0	2073	11	US-09-758-759-158	Sequence 158, App
C 38	35.4	5.0	96587	11	US-09-997-722-250	Sequence 250, App
C 39	35.4	5.0	109519	11	US-09-758-759-1	Sequence 1, App1
C 40	35.2	5.0	4941	13	US-10-001-189-53	Sequence 53, App1
C 41	35.2	5.0	4951	13	US-10-001-189-51	Sequence 51, App1
C 42	35.2	5.0	4952	13	US-10-001-189-52	Sequence 52, App1
C 43	35	5.0	51	9	US-09-969-652-14	Sequence 14, App1
C 44	34.6	4.9	330	18	US-10-425-115-68281	Sequence 68281, A
C 45	34.4	4.9	887	16	US-10-424-599-40607	Sequence 40607, A

ALIGNMENTS

RESULT 1
US-10-001-189-40
Sequence 40, Application US/10001189
Publication No. US20020173634A1
GENERAL INFORMATION:
APPLICANT: FRASER JR., MALCOLM J.
APPLICANT: LI, XU
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
FILE REFERENCE: 835910-92098
CURRENT APPLICATION NUMBER: US/10/001,189
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,984
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,677
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 707
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge
OTHER INFORMATION: sequence
US-10-001-189-40
Query Match 100.0%; Score 707; DB 13; Length 707;
Best Local Similarity 100.0%; Pred. No. 6,3e-236;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGATCCATGCGTCATTTTACGACGACTTCTTACGAGTTATCTAGTCATGAGG 60
|||||

```

Db      1  GGATCCCATGCGTCATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Qy      61  ATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCAAAGCGGCGCTATCTGGCA 120
Db      61  ATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCAAAGCGGCGCTATCTGGCA 120
Qy      121  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCAATGGAAGATTG 180
Db      121  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCAATGGAAGATTG 180
Qy      181  CCGAGATGACTGCTGCTGCAATGACGTTGACGAAACGACGCTTAACCATATATTC 240
Db      181  CCGAGATGACTGCTGCTGCAATGACGTTGACGAAACGACGCTTAACCATATATTC 240
Qy      241  GGAAGGTGTGCCATGACAGCCCTTTAAAGGTAACTGTTGACAGCCGCACTGGGATA 300
Db      241  GGAAGGTGTGCCATGACAGCCCTTTAAAGGTAACTGTTGACAGCCGCACTGGGATA 300
Qy      301  CCAATTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACGCGGAGCCATCAGCA 360
Db      301  CCAATTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACGCGGAGCCATCAGCA 360
Qy      361  ACCGGAACAATACCGGAGCAGCCGGAACCTGCGGTGCAATTAATGACAGC 420
Db      361  ACCGGAACAATACCGGAGCAGCCGGAACCTGCGGTGCAATTAATGACAGC 420
Qy      421  GTGCGGCTGGGATTAATGACGAGGAGGAGTCTGCTGAGTGTGAGTGTGAGTGTG 480
Db      421  GTGCGGCTGGGATTAATGACGAGGAGGAGTCTGCTGAGTGTGAGTGTGAGTGTG 480
Qy      481  GGCATGATGATACCCGCTGAGTTTACCGGCGGCGGCGCTGCTTACGCTTTTGAAC 540
Db      481  GGCATGATGATACCCGCTGAGTTTACCGGCGGCGGCGCTGCTTACGCTTTTGAAC 540
Qy      541  CCGTGAAGACGCGGAGACTCGCGGTGCAAAATGTTTACAGGTGAGGAGATG 600
Db      541  CCGTGAAGACGCGGAGACTCGCGGTGCAAAATGTTTACAGGTGAGGAGATG 600
Qy      601  AGATGCTGACACGCTGCAAGACGAGCTAGATTAACCTTAAGAAATATCATATTG 660
Db      601  AGATGCTGACACGCTGCAAGACGAGCTAGATTAACCTTAAGAAATATCATATTG 660
Qy      661  TGAAGTACGTTAAAGATATCATGCTTAATTAATGACCATGAGATCC 707
Db      661  TGAAGTACGTTAAAGATATCATGCTTAATTAATGACCATGAGATCC 707

```

RESULT 2

```

/ Sequence 46, Application US/10001189
/ Publication No. US20020173634A1
/ GENERAL INFORMATION:
/ APPLICANT: FRASER JR., MALCOLM J.
/ APPLICANT: LI, XU
/ TITLE OF INVENTION: BEAM, TERESA
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
/ TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
/ FILE REFERENCE: 835910-92098
/ CURRENT APPLICATION NUMBER: US/10/001,189
/ PRIOR FILING DATE: 2001-10-30
/ PRIOR APPLICATION NUMBER: 60/244,984
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 60/244,677
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 4613
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: pCRIT-ITR

```

```

/ OTHER INFORMATION: sequence
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (344)..(922)
US-10-001-189-46

```

```

Query Match      100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred. No. 1,7e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  GGATCCCATGCGTCATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db      294  GGATCCCATGCGTCATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353
Qy      61  ATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCAAAGCGGCGCTATCTGGCA 120
Db      354  ATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCAAAGCGGCGCTATCTGGCA 413
Qy      121  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCAATGGAAGATTG 180
Db      414  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCAATGGAAGATTG 473
Qy      181  CCGAGATGACTGCTGCTGCAATGACGTTGACGAAACGACGCTTAACCATATATTC 240
Db      474  CCGAGATGACTGCTGCTGCAATGACGTTGACGAAACGACGCTTAACCATATATTC 533
Qy      241  GGAAGGTGTGCCATGACAGCCCTTTAAAGGTAACTGTTGACAGCCGCACTGGGATA 300
Db      534  GGAAGGTGTGCCATGACAGCCCTTTAAAGGTAACTGTTGACAGCCGCACTGGGATA 593
Qy      301  CCAATTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACGCGGAGCCATCAGCA 360
Db      594  CCAATTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACGCGGAGCCATCAGCA 653
Qy      361  ACCGGAACAATACCGGAGCAGCCGGAACCTGCGGTGCAATTAATGACAGC 420
Db      654  ACCGGAACAATACCGGAGCAGCCGGAACCTGCGGTGCAATTAATGACAGC 713
Qy      421  GTGCGGCTGGGATTAATGACGAGGAGGAGTCTGCTGAGTGTGAGTGTGAGTGTG 480
Db      714  GTGCGGCTGGGATTAATGACGAGGAGGAGTCTGCTGAGTGTGAGTGTGAGTGTG 773
Qy      481  GGCATGATGATACCCGCTGAGTTTACCGGCGGCGGCGCTGCTTACGCTTTTGAAC 540
Db      774  GGCATGATGATACCCGCTGAGTTTACCGGCGGCGGCGCTGCTTACGCTTTTGAAC 833
Qy      541  CCGTGAAGACGCGGAGACTCGCGGTGCAAAATGTTTACAGGTGAGGAGATG 600
Db      834  CCGTGAAGACGCGGAGACTCGCGGTGCAAAATGTTTACAGGTGAGGAGATG 893
Qy      601  AGATGCTGACACGCTGCAAGACGAGCTAGATTAACCTTAAGAAATATCATATTG 660
Db      894  AGATGCTGACACGCTGCAAGACGAGCTAGATTAACCTTAAGAAATATCATATTG 953
Qy      661  TGAAGTACGTTAAAGATATCATGCTTAATTAATGACCATGAGATCC 707
Db      954  TGAAGTACGTTAAAGATATCATGCTTAATTAATGACCATGAGATCC 1000

```

RESULT 3

```

/ Sequence 48, Application US/10001189
/ Publication No. US20020173634A1
/ GENERAL INFORMATION:
/ APPLICANT: FRASER JR., MALCOLM J.
/ APPLICANT: LI, XU
/ TITLE OF INVENTION: BEAM, TERESA
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
/ TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
/ FILE REFERENCE: 835910-92098
/ CURRENT APPLICATION NUMBER: US/10/001,189
/ CURRENT FILING DATE: 2001-10-30

```

PRIOR APPLICATION NUMBER: 60/244,984
 PRIOR FILING DATE: 2000-11-01
 PRIOR APPLICATION NUMBER: 60/244,677
 PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 70
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 48
 LENGTH: 8999
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EYFP
 OTHER INFORMATION: sequence
 US-10-001-189-48

Query Match 100.0%; Score 707; DB 13; Length 8999;
 Best Local Similarity 100.0%; Pred. No. 2,4e-235;
 Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATCCATGCGTCAATTTTACGACATATCTTTCTAGGGTTAATAGTCATCAGG 60
 DB 8950 GGATCCCATGCGTCAATTTTACGACATATCTTTCTAGGGTTAATAGTCATCAGG 8891
 QY 61 ATCATATGTCGGGCTTTTCCGGCTCAGTCATCCGCAAGCTGGGCTATCTGGGCA 120
 DB 8890 ATCATATGTCGGGCTTTTCCGGCTCAGTCATCCGCAAGCTGGGCTATCTGGGCA 8831
 QY 121 TCGGGGAGGAAGAACCCCGTCCCTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
 DB 8830 TCGGGGAGGAAGAACCCCGTCCCTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 8771
 QY 181 CCGAGGATGACTGCTGCTGCACTTGAAGCGAAACGACGTTTACATGATGATTC 240
 DB 8770 CCGAGGATGACTGCTGCTGCACTTGAAGCGAAACGACGTTTACATGATGATTC 8711
 QY 241 GGGAAAGGTGGCCATGACGCTTTTAAAGGTAAGCTTTTCAAGGCCACTGGGATA 300
 DB 8710 GGGAAAGGTGGCCATGACGCTTTTAAAGGTAAGCTTTTCAAGGCCACTGGGATA 8651
 QY 301 CCAATTGTCGGGCTTTTCCGGAACAAGTCCGATGCTGACGCCAAGCGCATCAGCA 360
 DB 8650 CCAATTGTCGGGCTTTTCCGGAACAAGTCCGATGCTGACGCCAAGCGCATCAGCA 8591
 QY 361 ACCGGAACATATCCGGGCAAGCGGAACTGCGCGGCTGTCAGATTAATGACAGCG 420
 DB 8590 ACCGGAACATATCCGGGCAAGCGGAACTGCGCGGCTGTCAGATTAATGACAGCG 8531
 QY 421 GTGGGGGCTGGGATTTAGTCAAGCGAGGAGGATTCCTGGCTGGATGCGGAGAAAT 480
 DB 8530 GTGGGGGCTGGGATTTAGTCAAGCGAGGAGGATTCCTGGCTGGATGCGGAGAAAT 8471
 QY 481 GGACATGATATCCCGGATGATTACCCGCGGCGCGCTCGTTCACTTCACTTTTGAAC 540
 DB 8470 GGACATGATATCCCGGATGATTACCCGCGGCGCGCTCGTTCACTTCACTTTTGAAC 8411
 QY 541 CCGTGAAGAGCGGCAAGCTCGCGGTGCAAAATGTTTTCACAGCTGATGAGCAATGA 600
 DB 8410 CCGTGAAGAGCGGCAAGCTCGCGGTGCAAAATGTTTTCACAGCTGATGAGCAATGA 8351
 QY 601 AGATGCTGACAGGCTGACAGAACGACGAGTATTAACCTTAGAAAGATATCATATTG 660
 DB 8350 AGATGCTGACAGGCTGACAGAACGACGAGTATTAACCTTAGAAAGATATCATATTG 8291
 QY 661 TGAAGTACGTTAAAGATATCATATCGTAAATTAAGCGATGGATCC 707
 DB 8290 TGAAGTACGTTAAAGATATCATATCGTAAATTAAGCGATGGATCC 8244

RESULT 4
 US-10-001-189-49/c
 Sequence 49, Application US/10001189
 Publication No. US20020173634A1
 GENERAL INFORMATION:

APPLICANT: FRASER JR., MALCOLM J.
 APPLICANT: LI, XU
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
 TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 TITLE OF INVENTION: VECTOR PIGGYBAC
 FILE REFERENCE: 835910-92098
 CURRENT APPLICATION NUMBER: US/10/001,189
 PRIOR FILING DATE: 2001-10-30
 PRIOR APPLICATION NUMBER: 60/244,984
 PRIOR FILING DATE: 2000-11-01
 PRIOR APPLICATION NUMBER: 60/244,677
 PRIOR FILING DATE: 2000-10-31
 NUMBER OF SEQ ID NOS: 70
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 49
 LENGTH: 9012
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
 OTHER INFORMATION: sequence
 US-10-001-189-49

Query Match 100.0%; Score 707; DB 13; Length 9012;
 Best Local Similarity 100.0%; Pred. No. 2,4e-235;
 Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATCCATGCGTCAATTTTACGACATATCTTTCTAGGGTTAATAGTCATCAGG 60
 DB 8963 GGATCCCATGCGTCAATTTTACGACATATCTTTCTAGGGTTAATAGTCATCAGG 8904
 QY 61 ATCATATGTCGGGCTTTTCCGGCTCAGTCATCCGCAAGCTGGGCTATCTGGGCA 120
 DB 8903 ATCATATGTCGGGCTTTTCCGGCTCAGTCATCCGCAAGCTGGGCTATCTGGGCA 8844
 QY 121 TCGGGGAGGAAGAACCCCGTCCCTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
 DB 8843 TCGGGGAGGAAGAACCCCGTCCCTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 8784
 QY 181 CCGAGGATGACTGCTGCTGCACTTGAAGCGAAACGACGTTTACATGATGATTC 240
 DB 8783 CCGAGGATGACTGCTGCTGCACTTGAAGCGAAACGACGTTTACATGATGATTC 8724
 QY 241 GGGAAAGGTGGCCATGACGCTTTTAAAGGTAAGCTTTTCAAGGCCACTGGGATA 300
 DB 8723 GGGAAAGGTGGCCATGACGCTTTTAAAGGTAAGCTTTTCAAGGCCACTGGGATA 8664
 QY 301 CCAATTGTCGGGCTTTTCCGGAACAAGTCCGATGCTGACGCCAAGCGCATCAGCA 360
 DB 8663 CCAATTGTCGGGCTTTTCCGGAACAAGTCCGATGCTGACGCCAAGCGCATCAGCA 8604
 QY 361 ACCGGAACATATCCGGGCAAGCGGAACTGCGCGGCTGTCAGATTAATGACAGCG 420
 DB 8603 ACCGGAACATATCCGGGCAAGCGGAACTGCGCGGCTGTCAGATTAATGACAGCG 8544
 QY 421 GTGGGGGCTGGGATTTAGTCAAGCGAGGAGGATTCCTGGCTGGATGCGGAGAAAT 480
 DB 8543 GTGGGGGCTGGGATTTAGTCAAGCGAGGAGGATTCCTGGCTGGATGCGGAGAAAT 8484
 QY 481 GGACATGATATCCCGGATGATTACCCGCGGCGCGCTCGTTCACTTCACTTTTGAAC 540
 DB 8483 GGACATGATATCCCGGATGATTACCCGCGGCGCGCTCGTTCACTTCACTTTTGAAC 8424
 QY 541 CCGTGAAGAGCGGCAAGCTCGCGGTGCAAAATGTTTTCACAGCTGATGAGCAATGA 600
 DB 8423 CCGTGAAGAGCGGCAAGCTCGCGGTGCAAAATGTTTTCACAGCTGATGAGCAATGA 8364
 QY 601 AGATGCTGACAGGCTGACAGAACGACGAGTATTAACCTTAGAAAGATATCATATTG 660
 DB 8363 AGATGCTGACAGGCTGACAGAACGACGAGTATTAACCTTAGAAAGATATCATATTG 8304
 QY 661 TGAAGTACGTTAAAGATATCATATCGTAAATTAAGCGATGGATCC 707

Db 8303 TGACGTACGTTAAAGATATCATGCTGTAATAATTGACCATGGATCC 8257

RESULT 5
US-10-001-189-50/c
Sequence 50, Application US/10001189
Publication No. US20020173634A1
GENERAL INFORMATION:
APPLICANT: FRASER JR., MALCOLM J.
APPLICANT: LI, XU
APPLICANT: BEAM, TERESA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
FILE REFERENCE: 835910-92098
CURRENT APPLICATION NUMBER: US/10/001,189
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,984
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,677
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 50
LENGTH: 9013
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EGFP
US-10-001-189-50

Query Match
Best Local Similarity 100.0%; Score 707; DB 13; Length 9013;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGCGTCAATTTTACGACAGATCTTTCTTGGGTTATCTAGTGCATAGG 60
Db 8964 GGATCCCATGCGTCAATTTTACGACAGATCTTTCTTGGGTTATCTAGTGCATAGG 60
Qy 61 ATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGCGCTATCTGGCA 120
Db 8904 ATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGCGCTATCTGGCA 120
Qy 121 TCGGGAGGAGAAAGCCCGTCCTTTTCCCGAGTTGAAGGGCATGGAAAGATTG 180
Db 8844 TCGGGAGGAGAAAGCCCGTCCTTTTCCCGAGTTGAAGGGCATGGAAAGATTG 180
Qy 181 CCGAGATGACTGCTGCTGATGACGTTGAGGAAAGCAAGTTTACATGATGATTC 240
Db 8784 CCGAGATGACTGCTGCTGATGACGTTGAGGAAAGCAAGTTTACATGATGATTC 240
Qy 241 GGGAAAGGTGTCGATGACGCTTTTACGTTGACGTTGACGTTGACGTTGACGTTG 300
Db 8724 GGGAAAGGTGTCGATGACGCTTTTACGTTGACGTTGACGTTGACGTTGACGTTG 300
Qy 301 CCAATTGTCGCGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATGATG 360
Db 8664 CCAATTGTCGCGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATGATG 360
Qy 361 ACCGGAACAATATCCGCGACAGCCGGAACCTGCGGTGATGATGATGATGATGATG 420
Db 8604 ACCGGAACAATATCCGCGACAGCCGGAACCTGCGGTGATGATGATGATGATGATG 420
Qy 421 GTCCGCGCTGGGATATTAATGTCAGCGAGGACGGGATCTGCTGATGATGATGATG 480
Db 8544 GTCCGCGCTGGGATATTAATGTCAGCGAGGACGGGATCTGCTGATGATGATGATG 480
Qy 481 GGACATGATATCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 8484 GGACATGATATCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 540

Qy 541 CCGTGGAGGACGGGACAGCTGCGGTGCAATGCTTTTACAGCTGATGAGACAGATGA 600
Db 8424 CCGTGGAGGACGGGACAGCTGCGGTGCAATGCTTTTACAGCTGATGAGACAGATGA 600
Qy 601 AGATGCTGACACCGTGGAGAAACGCAAGCTGATTAACCTTGAAGATTAATGATG 660
Db 8364 AGATGCTGACACCGTGGAGAAACGCAAGCTGATTAACCTTGAAGATTAATGATG 660
Qy 661 TGACGTACGTTAAAGATATCATGCTGTAATAATTGACCATGGATCC 707
Db 8304 TGACGTACGTTAAAGATATCATGCTGTAATAATTGACCATGGATCC 8258

RESULT 6
US-10-001-189-54
Sequence 54, Application US/10001189
Publication No. US20020173634A1
GENERAL INFORMATION:
APPLICANT: FRASER JR., MALCOLM J.
APPLICANT: LI, XU
APPLICANT: BEAM, TERESA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
FILE REFERENCE: 835910-92098
CURRENT APPLICATION NUMBER: US/10/001,189
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,984
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,677
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 54
LENGTH: 4943
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: pBS-JTR-BCFP
US-10-001-189-54

Query Match
Best Local Similarity 96.9%; Score 685; DB 13; Length 4943;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATGCGTCAATTTTACGACAGATCTTTCTTGGGTTATCTAGTGCATAGG 60
Db 719 GGATCCCATGCGTCAATTTTACGACAGATCTTTCTTGGGTTATCTAGTGCATAGG 60
Qy 61 ATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGCGCTATCTGGCA 120
Db 779 ATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGCGCTATCTGGCA 120
Qy 121 TCGGGAGGAGAAAGCCCGTCCTTTTCCCGAGTTGAAGGGCATGGAAAGATTG 180
Db 839 TCGGGAGGAGAAAGCCCGTCCTTTTCCCGAGTTGAAGGGCATGGAAAGATTG 180
Qy 181 CCGAGATGACTGCTGCTGATGACGTTGAGGAAAGCAAGTTTACATGATGATTC 240
Db 899 CCGAGATGACTGCTGCTGATGACGTTGAGGAAAGCAAGTTTACATGATGATTC 240
Qy 241 GGGAAAGGTGTCGATGACGCTTTTACGTTGACGTTGACGTTGACGTTGACGTTG 300
Db 959 GGGAAAGGTGTCGATGACGCTTTTACGTTGACGTTGACGTTGACGTTGACGTTG 300
Qy 301 CCAATTGTCGCGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATGATGATG 360
Db 1019 CCAATTGTCGCGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATGATGATG 360
Qy 361 ACCGGAACAATATCCGCGACAGCCGGAACCTGCGGTGATGATGATGATGATGATG 420
Db 1079 ACCGGAACAATATCCGCGACAGCCGGAACCTGCGGTGATGATGATGATGATGATG 420

QY	421	GTGGGGCGCTGGGATATTAAGTCAGACGAGACGGGTATCTCGCTGATATGCCGAGAAAT	480
Db	1139	GTGGGGCGCTGGGATATTAAGTCAGACGAGACGGGTATCTCGCTGATATGCCGAGAAAT	1198
QY	481	GGACATGAGTACCCCGGTAGTTATCCCGGCGGGCGCTCGTTCACTTACAGCTTTTGAAC	540
Db	1199	GGACATGAGTACCCCGGTAGTTATCCCGGCGG-----CTCGTTCACTTACAGCTTTTGAAC	1252
QY	541	CCGTGAGAGACGGGAGACTCGCGGTGCAATTCGTATTAACAAGCTGATGAGACAGATGA	600
Db	1253	CCGTGAGAGACGGGAGACTCGCGGTGCAATTCGTATTAACAAGCTGATGAGACAGATGA	1312
QY	601	AGATGCTTCGACACGCTGCAGAACACGACGTAGATTTAACCTTAGAAGATATCATATTG	660
Db	1313	AGATGCTTCGACACGCTGCAGAACACGACGTAGATTTAACCTTAGAAGATATCATATTG	1372
QY	661	TGACGTACGTTAAAGATATATCATGCTTAAATTGACGATGGGATCC	707
Db	1373	TGACGTACGTTAAAGATATATCATGCTTAAATTGACGATGGGATCC	1419

RESULT 7

US-10-001-189-55
Sequence 55, Application US/10001189
Publication NO. US20020173634A1
GENERAL INFORMATION:
APPLICANT: FRASER JR., MALCOLM J.
APPLICANT: LI, XU
APPLICANT: BEAM, TERESA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSCRIPTION USING
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSCRIPTION
FILE REFERENCE: 835910-92098
CURRENT APPLICATION NUMBER: US/10/001,189
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,984
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/244,677
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 55
LENGTH: 4944
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
US-10-001-189-55

Query Match	96.9%	Score 685	DB 13	Length 4944
Best Local Similarity	99.2%	Pred. No. 8.6e-228		
Matches 701	Conservative 0	Mismatches 0	Indels 6	Gaps 1

Oy	1	GGATCCCATGGGTGAATTTAAGGAGCATATTTCTTAAAGGTTATCTAAGTCGATCAGG	60
Db	719	GGATCCCATGGGTCAATTTTAAAGCAAGCATATTTCTTAAAGGTTATCTAAGTCGATCAGG	778
Oy	61	ATCATATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA	120
Db	779	ATCATATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA	838
Oy	121	TCGGGGAGAGAAAGCCCGTCCTTTTCCCGGAGGTTGAAGCGGCGATGGAAAGAGTTTG	180
Db	839	TCGGGGAGAGAAAGCCCGTCCTTTTCCCGGAGGTTGAAGCGGCGATGGAAAGAGTTTG	898
Oy	181	CCGAGGAGATGACTGCTGCTGCATTAAGTTTGAAGGAAAACGACCGTTTACCATGATGATTC	240
Db	899	CCGAGGAGATGACTGCTGCTGCATTAAGTTTGAAGGAAAACGACCGTTTACCATGATGATTC	958
Oy	241	GGGAAAGGTGTGGCCATGACGCGCTTTAACGTTGAACGTGTTCGTTCAAGGCAACTTGGGATA	300

Db	959	GGGAAAGGTGTGGCCATGCACGCTTTTAAACGGTAGACTGTTCGTTACAGCCACCTGGGATA	1018
Qy	301	CCAGTCTGTGCGGGCTTTTTCGGGACACAGTTTCCGGATGTCAGCCCGAAGCGCATCGCA	360
Db	1019	CCAATTGGTGGCGGCTTTTCCGGACACAGTTTCCGGATGTCAGCCCGAAGCGCATCGCA	1078
Qy	361	ACCCGAACAATACCGGCGACAGCCGGAACTGCGTGCCGGGTGTGCAGATTAAATGACACGC	420
Db	1079	ACCCGAACAATACCGGCGACAGCCGGAACTGCGTGCCGGGTGTGCAGATTAAATGACACGC	1138
Qy	421	GTGCGGGGCTGTGGATTATTAAGTCAGCGGACAGGAGTCTCTGGCTGGATGCCGACGAAT	480
Db	1139	GTGCGGGGCTGTGGATTATTAAGTCAGCGGACAGGAGTCTCTGGCTGGATGCCGACGAAT	1198
Qy	481	GGACATGGAATACCCCGTGAAGTTACCCGGCGGGCGCGCTGTTCATTCAAGTTTTCAAAC	540
Db	1199	GGACATGGAATACCCCGTGAAGTTACCCGGCGG-----CTGTTTCATTCAAGTTTTCAAAC	1252
Qy	541	CCGTGTGAGACGGGCGACACTCGCGGTGCAAAATGTGTTTACAGCGTATATGAGCGAATGA	600
Db	1253	CCGTGTGAGACGGGCGACACTCGCGGTGCAAAATGTGTTTACAGCGTATATGAGCGAATGA	1312
Qy	601	AGATGCTCGACAGCGTGCAGAACCGAGCTAGATTAAACCTTGAAGATTAATCATATTG	660
Db	1313	AGATGCTCGACAGCGTGCAGAACCGAGCTAGATTAAACCTTGAAGATTAATCATATTG	1372
Qy	661	TGACGTAACGTTAAAGATTAATCATATCGTAAATTAAGCATGGGATTC	707
Db	1373	TGACGTAACGTTAAAGATTAATCATATCGTAAATTAAGCATGGGATTC	1419

RESULT 8

```

US-10-001-189-56
: Sequence 56, Application US/10001189
: Publication No. US20020173534A1
: GENERAL INFORMATION:
: APPLICANT: FRASER JR., MALCOLM J.
: APPLICANT: LI, XU
: APPLICANT: BEAM, TERESA
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
: TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
: TITLE OF INVENTION: VECTOR PIGGYBAC
: FILE REFERENCE: 835910-92098
: CURRENT APPLICATION NUMBER: US/10/001,189
: CURRENT FILING DATE: 2001-10-30
: PRIOR APPLICATION NUMBER: 60/244,984
: PRIOR FILING DATE: 2000-11-01
: PRIOR APPLICATION NUMBER: 60/244,677
: PRIOR FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 70
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 56
:
: LENGTH: 4944
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: pBS-ITR-EYFP
: OTHER INFORMATION: sequence
US-10-001-189-56

```

Query Match	96.9%	Score 685	DB 13	Length 4944
Best Local Similarity	99.2%	Pred. No. 8	6e-228	
Matches 701	Conservative	0	Mismatches	0
			Indels	6
			Gaps	1

QY	1	GGATCCCATGCGTCAATTTTATGCGAGACTATCTTTCTAGAGGTAAATCTAGCTGATCAGG	60
Db	719	GGATCCCATGCGTCAATTTTATGCGAGACTATCTTTCTAGAGGTAAATCTAGCTGATCAGG <td>778</td>	778
QY	61	ATCATATTCGTGGGCTCTTTTTTCCGGCTCACTCATGCCCAAGCTGGCGCATCTATCTGGACA	120
Db	779	ATCATATTCGTGGGCTCTTTTTTCCGGCTCACTCATGCCCAAGCTGGCGCATCTATCTGGACA	838
QY	121	TGGGAGGAGAAAGCCCGTGCCTTTTCCCGCGAGGTGGAACGCGCATATGAAAGATTTG	180

REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 46819
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-114-170-72

Query Match 80.1%; Score 566.2; DB 14; Length 46819;
Best Local Similarity 97.8%; Pred. No. 9.7e-186;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

45 ATCTAGCTGCATCGAGATCATATGTCGGGCTTTTCCGGCTCAGTCATCGCCCAAGC 104
14795 ATCCAGCTGCATCGAGATCATATGTCGGGCTTTTCCGGCTCAGTCATCGCCCAAGC 14854
105 TGCGGCTATCTGGGCTATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGTTGAAGCG 164
14855 TGCGGCTATCTGGGCTATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGTTGAAGCG 14914
165 GCATGGAAGAAGTTTCCGAGAGATGACTGCTGCTGATGACGTTGAGCGAAAACGACG 224
14915 GCATGGAAGAAGTTTCCGAGAGATGACTGCTGCTGATGACGTTGAGCGAAAACGACG 14974
225 TTTCACATGATGATTCGGGAGAGTGTGCGCATGACACGCGCTTAAACGTTGACCTGCTT 284
14975 TTTCACATGATGATTCGGGAGAGTGTGCGCATGACACGCGCTTAAACGTTGACCTGCTT 15034
285 CAGGCCAATCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGATGACG 344
15035 CAGGCCAATCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGATGACG 15094
345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGAACTGCGCGGTGTG 404
15095 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGAACTGCGCGGTGTG 15154
405 CAGATTATGACACCGGTCGGGCTGAGATTAATGTCACGAGAGAGAGGATATCTGTCG 464
15155 CAGATTATGACACCGGTCGGGCTGAGATTAATGTCACGAGAGAGAGGATATCTGTCG 15214
465 TGATGCGCGCAGAAATGACATGATACCCCGTGAATTAACCGGCGGCGCGCTTCGTTG 524
15215 TGATGCGCGCAGAAATGACATGATACCCCGTGAATTAACCGGCGGCGCGCTTCGTTG 15274
525 ATTACGCTTTTGAACCGGTGAGAGACGGGAGACTCGCGGTGCAATGTTTTCACGC 584
15275 ATTACGCTTTTGAACCGGTGAGAGACGGGAGACTCGCGGTGCAATGTTTTCACGC 15334
585 GTGATGAGCAGATGAAGATGCTGACACGCTGCAAGAACAGCGAGCT 631
15335 GTGATGAGCAGATGAAGATGCTGACACGCTGCAAGAACAGCGAGCT 15381

RESULT 11
US-10-114-170-251/c
Sequence 251, Application US/10114170
Publication No. US20030023075A1
GENERAL INFORMATION:
APPLICANT: Blatner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady
STREET: 1 South Plinkney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 251:
SEQUENCE CHARACTERISTICS:
LENGTH: 22306
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-10-114-170-251

Query Match 79.4%; Score 561.4; DB 14; Length 22306;
Best Local Similarity 97.3%; Pred. No. 3.1e-184;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

45 ATCTAGCTGCATCGAGATCATATGTCGGGCTTTTCCGGCTCAGTCATCGCCCAAGC 104
17193 ATCTAGCTGCATCGAGATCATATGTCGGGCTTTTCCGGCTCAGTCATCGCCCAAGC 17134
105 TGCGGCTATCTGGGCTATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGTTGAAGCG 164
17133 TGCGGCTATCTGGGCTATCGGGAGAGAAAGCCCGTCCCTTTCCCGGAGTTGAAGCG 17074
165 GCATGGAAGAAGTTTCCGAGAGATGACTGCTGCTGATGACGTTGAGCGAAAACGACG 224
17073 GCATGGAAGAAGTTTCCGAGAGATGACTGCTGCTGATGACGTTGAGCGAAAACGACG 17014
225 TTTCACATGATGATTCGGGAGAGTGTGCGCATGACACGCGCTTAAACGTTGACCTGCTT 284
17013 TTTCACATGATGATTCGGGAGAGTGTGCGCATGACACGCGCTTAAACGTTGACCTGCTT 16954
285 CAGGCCAATCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGATGACG 344
16953 CAGGCCAATCTGGGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGATGACG 16894
345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGAACTGCGCGGTGTG 404
16893 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGAACTGCGCGGTGTG 16834
405 CAGATTATGACACCGGTCGGGCTGAGATTAATGTCACGAGAGAGAGGATATCTGTCG 464
16833 CAGATTATGACACCGGTCGGGCTGAGATTAATGTCACGAGAGAGAGGATATCTGTCG 16774
465 TGATGCGCGCAGAAATGACATGATACCCCGTGAATTAACCGGCGGCGCGCTTCGTTG 524
16773 TGATGCGCGCAGAAATGACATGATACCCCGTGAATTAACCGGCGGCGCGCTTCGTTG 16714

QY	525	ATTACGCTTTTGAACCCGTGGAGAGCGGACAGCTCGCGTGCAAATGTTTACAGC	584
Db	16713	ATTACGCTTTTGAACCCGTGGAGAGCGGACAGCCGCGGTGCAAATGTTTACAGC	
QY	585	GTGATGGAGCAGATGAGATGCTGCACGCTGCAGACACCGTAGCT	631
Db	16653	GTGATGGAGCAGATGAGATGCTGCACGCTGCAGACACCGTAGCT	

```

RESULT 12
US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 41
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXl-Bac
US-10-001-189-41

```

```

0y      482  GACATGATACCCCGTAGTTACCCGCGGCGCGCCTCGTTACATCAGCTT  533
Db      1463  GACATGATACCCCGTAGTTACCCGCGGCGCGCCTGGCGTAAATCATGCT  1514

RESULT 13
US-10-001-189-53/c
/ Sequence 53, Application US/10001189
/ Publication No. US20020173634A1
/ GENERAL INFORMATION:
/ APPLICANT: FRASER JR., MALCOLM J.
/ APPLICANT: LI, XU
/ APPLICANT: BEAM, TERESA
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
/ TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
/ TITLE OF INVENTION: VECTOR PIGGYBAC
/ FILE REFERENCE: 835910-92098
/ CURRENT APPLICATION NUMBER: US/10/001,189
/ CURRENT FILING DATE: 2001-10-30
/ PRIOR APPLICATION NUMBER: 60/244,984
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 60/244,677
/ PRIOR FILING DATE: 2000-10-31
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 53
/ LENGTH: 4841
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECFP
/ US-10-001-189-53

```

Db 649 GGATACCCCGTAGATTACCGCGCGCGCGC 619

RESULT 14
US-10-001-189-51/c

; Sequence 51, Application US/10001189
; Publication No. US20020173634A1

; GENERAL INFORMATION:

; APPLICANT: FRASER JR., MALCOLM J.

; APPLICANT: LI, XU

; APPLICANT: BEAM, TERESA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

; FILE REFERENCE: 835910-92098

; CURRENT APPLICATION NUMBER: US/10/001,189

; PRIOR FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/244,984

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 60/244,677

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 51

; LENGTH: 4951

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-BXP

; US-10-001-189-51

Query Match 72.3%; Score 511; DB 13; Length 4951;

Best Local Similarity 100.0%; Pred. No. 5.9e-167;

Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 CATCGTCAATTTTACGAGACTATCTTTAGGGTTATCTAGCTGATCATGATCATTA 66

Db 1129 CATCGTCAATTTTACGAGACTATCTTTAGGGTTATCTAGCTGATCATGATCATTA 1070

Qy 67 TCCTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 126

Db 1069 TCCTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 1010

Qy 127 AGAAGAAAGCCCGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 186

Db 1009 AGAAGAAAGCCCGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 950

Qy 187 ATGACTGCTGCTGATGAGCGTTGAGGAAACGACGTTTACCATGATGATGAGGAG 246

Db 949 ATGACTGCTGCTGATGAGCGTTGAGGAAACGACGTTTACCATGATGATGAGGAG 890

Qy 247 GTGTGGCCATGACGCTTTTACGAGTGAATGTTGTTCAAGGCCACCTGGGATACAGTT 306

Db 889 GTGTGGCCATGACGCTTTTACGAGTGAATGTTGTTCAAGGCCACCTGGGATACAGTT 830

Qy 307 CGTGGCGGCTTTTTCGGGCAAGTTCGGGATGTCAGCCGCAAGCGCATACGCAACCGCA 366

Db 829 CGTGGCGGCTTTTTCGGGCAAGTTCGGGATGTCAGCCGCAAGCGCATACGCAACCGCA 770

Qy 367 ACAATACCGGCGACGCGGAACTGCGGCGGATGTCAGTTAATGACGCGGTGCG 426

Db 769 ACAATACCGGCGACGCGGAACTGCGGCGGATGTCAGTTAATGACGCGGTGCG 710

Qy 427 CGCTGGGATATTAGTCAAGGAGACGCGGATTCCTGCTGATGTCGCGGAAATGACAT 486

Db 709 CGCTGGGATATTAGTCAAGGAGACGCGGATTCCTGCTGATGTCGCGGAAATGACAT 650

Qy 487 GGATACCCCGTAGATTACCGCGCGCGCGC 517

Db 649 GGATACCCCGTAGATTACCGCGCGCGCGC 619

RESULT 15
US-10-001-189-52/c

; Sequence 52, Application US/10001189
; Publication No. US20020173634A1

; GENERAL INFORMATION:

; APPLICANT: FRASER JR., MALCOLM J.

; APPLICANT: LI, XU

; APPLICANT: BEAM, TERESA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

; FILE REFERENCE: 835910-92098

; CURRENT APPLICATION NUMBER: US/10/001,189

; PRIOR FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/244,984

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 60/244,677

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 52

; LENGTH: 4952

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-BGFP

; US-10-001-189-52

Query Match 72.3%; Score 511; DB 13; Length 4952;

Best Local Similarity 100.0%; Pred. No. 5.9e-167;

Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 CATCGTCAATTTTACGAGACTATCTTTAGGGTTATCTAGCTGATCATGATCATTA 66

Db 1129 CATCGTCAATTTTACGAGACTATCTTTAGGGTTATCTAGCTGATCATGATCATTA 1070

Qy 67 TCCTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 126

Db 1069 TCCTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 1010

Qy 127 AGAAGAAAGCCCGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 186

Db 1009 AGAAGAAAGCCCGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGCTGCGCTATCTGGGCAATCGGG 950

Qy 187 ATGACTGCTGCTGATGAGCGTTGAGGAAACGACGTTTACCATGATGATGAGGAG 246

Db 949 ATGACTGCTGCTGATGAGCGTTGAGGAAACGACGTTTACCATGATGATGAGGAG 890

Qy 247 GTGTGGCCATGACGCTTTTACGAGTGAATGTTGTTCAAGGCCACCTGGGATACAGTT 306

Db 889 GTGTGGCCATGACGCTTTTACGAGTGAATGTTGTTCAAGGCCACCTGGGATACAGTT 830

Qy 307 CGTGGCGGCTTTTTCGGGCAAGTTCGGGATGTCAGCCGCAAGCGCATACGCAACCGCA 366

Db 829 CGTGGCGGCTTTTTCGGGCAAGTTCGGGATGTCAGCCGCAAGCGCATACGCAACCGCA 770

Qy 367 ACAATACCGGCGACGCGGAACTGCGGCGGATGTCAGTTAATGACGCGGTGCG 426

Db 769 ACAATACCGGCGACGCGGAACTGCGGCGGATGTCAGTTAATGACGCGGTGCG 710

Qy 427 CGCTGGGATATTAGTCAAGGAGACGCGGATTCCTGCTGATGTCGCGGAAATGACAT 486

Db 709 CGCTGGGATATTAGTCAAGGAGACGCGGATTCCTGCTGATGTCGCGGAAATGACAT 650

Qy 487 GGATACCCCGTAGATTACCGCGCGCGCGC 517

Db 649 GGATACCCCGTAGATTACCGCGCGCGCGC 619

Search completed: January 26, 2005, 17:51:21
Job time : 381.34 secs

This Page Blank (uspto)